

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 20:22:40 ; Search time 2633 Seconds

(without alignments)
12498.758 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032

Sequence: 1 ggcctcagagccagatgctct.....atagaagaaaaaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 16154066 seqs, 809774376 residues

Word size : 25

Total number of hits satisfying chosen parameters: 24

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	630	31.0	954	13	BI823850 603039012
2	581	28.6	593	14	BM969282 01-CF-ENO
3	551	27.1	571	13	BM129080 1f17c04.y
4	478	23.5	553	13	BM128370 1f13d12.x
5	478	23.5	553	13	BM128831 1f17c04.x
6	468	23.0	668	9	AL709927 DKP2p6860

7	318	15.6	680	12	BF878439	BF878439 MRO-ET010
8	288	14.2	362	10	AM002418	AM002418 wu61g03.x
9	283	13.9	419	10	AM572390	AM572390 xu02a06.x
10	150	7.4	417	10	AM572510	AM572510 xq16g11.x
11	136	6.7	358	12	BF056840	BF056840 7k10d12.x
12	136	6.7	358	12	BF197521	BF197521 7084a08.x
13	59	2.9	525	12	BF197521	BF197521 7084a08.x
14	59	2.9	620	9	A1824100	A1824100 w146c01.x
15	59	2.9	695	12	BE857538	BE857538 7g01a08.x
16	47	2.3	735	12	BE858652	BE858652 7g01a09.x
17	27	1.3	497	13	BM246681	BM246681 xc41b06.x
18	27	1.3	536	13	BM245312	BM245312 K0741E04-
19	27	1.3	695	13	BG964671	BG964671 K0722H07-
20	27	1.3	783	13	BG963298	BG963298 602831875
21	27	1.3	852	13	BG966340	BG966340 602832826
22	27	1.3	1923	11	AK009113	AK009113 Mus muscu
23	25	1.2	501	13	BM036473	BM036473 Iub107.x
24	25	1.2	682	17	CNS028KQ	AL189971 Tetraodon

ALIGNMENTS

RESULT 1
BI823850 954 bp mRNA linear EST 04-OCT-2001
DEFINITION 603039012F1 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5179826 5',
LOCUS BI823850
ACCESSION BI823850
VERSION BI823850.1 GI:15935400
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
TITLE 1 (bases 1 to 954)
JOURNAL NIH-MGC http://mgc.nci.nih.gov/
COMMENT National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LINT)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINT at:
http://image.llnl.gov
Plate: L1AM11448 row: h column: 03
High quality sequence stop: 856.
Location/Qualifiers

FEATURES

1. 954
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5179826"
/clone_lib="NIH_MGC_115"
/lab_host="DH10B"
/note="Organ: pooled brain, lung, testis; Vector: pcwv-sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research genetics tracking code 021. Note: this is a NIH_MGC Library."

BASE COUNT 199 a 297 c 250 g 208 t
ORIGIN
Query Match 31.0%; Score 630; DB 13; Length 954;
Best Local Similarity 100.0%; Pred. No. 1.1e-186;
Matches 630; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 160 CAAGGCTTCACCTGACACATGCTAGCTGCTAAATAAATGAAAGCTGCTGTTTCT 219
      |||
      |||
      |||
Db 56 CAAGGCTTCACCTGACACATGCTAGCTGCTAAATAAATGAAAGCTGCTGTTTCT 115
      |||
      |||
      |||
QY 220 GGTTCCTCCAGATGGCCATCTTGGCTCTATTCTTCCACATGTACAGCCACACATCAGCTC 279
      |||
      |||
      |||
Db 116 GGTTCCTCCAGATGGCCATCTTGGCTCTATTCTTCCACATGTACAGCCACACATCAGCTC 175
      |||
      |||
      |||
QY 280 CCGTGTATGAGGACAGCCGAGGCGATGACAGCGTGGTCTGCTTCCGGGCGCTC 339
      |||
      |||
      |||
Db 176 CCGTGTATGAGGACAGCCGAGGCGATGACAGCGTGGTCTGCTTCCGGGCGCTC 235
      |||
      |||
      |||
QY 340 TGGCTCTCTTTTGTGGGGGAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGA 399
      |||
      |||
      |||
Db 236 TGGCTCTCTTTTGTGGGGGAGCTTTTGGGAGCACCAGATGTTTCTACCTGATGA 295
      |||
      |||
      |||
QY 400 GCGCGCTGGCAGCTGTGATGTGACCTTCAAGCAGACAGCCGCTGATGCTGACATGGC 459
      |||
      |||
      |||
Db 296 GCGCGCTGGCAGCTGTGATGTGACCTTCAAGCAGACAGCCGCTGATGCTGACATGGC 355
      |||
      |||
      |||
QY 460 TGTGGGATCTGATAGCGGCGCTTCTTGTGCGACATGAGGCTTGTGATGCCATAC 519
      |||
      |||
      |||
Db 356 TGTGGGATCTGATAGCGGCGCTTCTTGTGCGACATGAGGCTTGTGATGCCATAC 415
      |||
      |||
      |||
QY 520 GGAACCTGGTCCCGGAGAGATGTCAGCCCTTTTCAAGTGGGAGAAAGCCGGGCGCTGTG 579
      |||
      |||
      |||
Db 416 GGAACCTGGTCCCGGAGAGATGTCAGCCCTTTTCAAGTGGGAGAAAGCCGGGCGCTGTG 475
      |||
      |||
      |||
QY 580 TTTCGACCTGCTGTGACATCATCCACAAATGAAATGATCCCGGCGCTACTGACG 639
      |||
      |||
      |||
Db 476 TTTCGACCTGCTGTGACATCATCCACAAATGAAATGATCCCGGCGCTACTGACG 535
      |||
      |||
      |||
QY 640 GCTTCGTGAGTGAACAGCCCTTTGAGGTGTGAGAGAGGCTGCCCTCTCTACAGCA 699
      |||
      |||
      |||
Db 536 GCTTCGTGAGTGAACAGCCCTTTGAGGTGTGAGAGAGGCTGCCCTCTCTACAGCA 595
      |||
      |||
      |||
QY 700 CGTGTGCTCAGAGAGGTGGCTTCTTCAACCTGACATGCCCTTACCGGCTGTGAAGA 759
      |||
      |||
      |||
Db 596 CGTGTGCTCAGAGAGGTGGCTTCTTCAACCTGACATGCCCTTACCGGCTGTGAAGA 655
      |||
      |||
      |||
QY 760 CCCCTCCCTCAACCTGCATATCGTGACCT 789
      |||
      |||
      |||
Db 656 CCCCTCCCTCAACCTGCATATCGTGACCT 685
      |||
      |||
      |||
RESULT 2
BM969292/c 593 bp mRNA linear EST 21-MAR-2002
LOCUS BM969292
DEFINITION UI-CF-ENO-acp-1-21-0-UI.s1 UI-CF-ENO Homo sapiens cDNA clone
VERSION BM969292
ACCESSION BM969292.1 GI:19586879
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 593)
AUTHORS Ronaldo M.F., Lennon G. and Soares M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
COMMENT Contact: McCray, PB
McCrays Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
cDNA library preparation: Dr. M. Bento Soares, University of Iowa

```

FEATURES

Source

CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Cloning Distribution: Researchers may obtain clones from Research
 Genetics (www.resgen.com).
 Seq primer: M13 FORWARD
 POLYA-Yes.

Location/Qualifiers

```

1..593
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="UI-CF-ENO-acp-1-21-0-UI"
/clone_lib="UI-CF-ENO"
/tissue_type="Primary Lung Cystic Fibrosis Epithelial
Cells"
/dev_stage="Adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: p773-Pac (Pharmacia) with a
modified polylinker; Site:1: Ecor I; Site:2: Not I;
UI-CF-ENO is a cDNA library containing the following
tissue(s): Primary Lung Cystic Fibrosis Epithelial Cells.
The library was constructed according to Ronaldo, Lennon
and Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an Ecor I adaptor, digested with Not I, and cloned
directionally into p773-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dr)18 tail. The sequence tag for this
library is CTCCTCAGGT.
TAG_Lib=UI-CF-ENO
TAG_Tissue=Human Lung Epithelial cell lines untreated LPS
6hr to LPS 24h
TAG_SEQ=CTCCTCAGGT"
BASE COUNT 157 a 127 c 137 g 172 t
ORIGIN

```

Query Match

Best Local Similarity 100.0%; Score 581; DB 14; Length 593;
 Matches 581; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1436 ACATGCTGTGGGTATACACTGATGTGAGTTGTGTCACACGTCCTCAAGCAGAGA 1495
      |||
      |||
      |||
Db 593 ACATGCTGTGGGTATACACTGATGTGAGTTGTGTCACACGTCCTCAAGCAGAGA 534
      |||
      |||
      |||
QY 1496 CTTTGTGTCATGCTGTGCTAGAAAACAGACAGTGGGAACTATGTGACAGACAT 1555
      |||
      |||
      |||
Db 533 CTTTGTGTCATGCTGTGCTAGAAAACAGACAGTGGGAACTATGTGACAGACAT 474
      |||
      |||
      |||
QY 1556 CCCACAGTGAACAGGGTATTGCTCTTCTTTTCTTGATCTTCTGCTGGGACAGC 1615
      |||
      |||
      |||
Db 473 CCCACAGTGAACAGGGTATTGCTCTTCTTTTCTTGATCTTCTGCTGGGACAGC 414
      |||
      |||
      |||
QY 1616 TTCAGAGCTTTGGGCGCTGGAGGCGCTATTAGCAGCAGACAGTACAGTGAATTGATC 1675
      |||
      |||
      |||
Db 413 TTCAGAGCTTTGGGCGCTGGAGGCGCTATTAGCAGCAGACAGTACAGTGAATTGATC 354
      |||
      |||
      |||
QY 1676 CATAAAGCTCCCTGTCCACATCTTCCCAATGGGGAATGGATCTTTCACAAAGAGCTCA 1735
      |||
      |||
      |||
Db 353 CATAAAGCTCCCTGTCCACATCTTCCCAATGGGGAATGGATCTTTCACAAAGAGCTCA 294
      |||
      |||
      |||
QY 1736 CCAGCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGGAGTTCCTCATGATTAAGA 1795
      |||
      |||
      |||
Db 293 CCAGCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGGAGTTCCTCATGATTAAGA 234
      |||
      |||
      |||
QY 1796 AGGAAGTGGAAAGAGTGTGATGCTTATGAGCTTGAACATACAGATATGCTTA 1855
      |||
      |||
      |||
Db 233 AGGAAGTGGAAAGAGTGTGATGCTTATGAGCTTGAACATACAGATATGCTTA 174
      |||
      |||
      |||
QY 1856 TCAGAAATATGAACAAAATCTCTGACAAAAGACAAAGCTTATAGTTACAGGGTGC 1915
      |||
      |||
      |||
Db 173 TCAGAAATATGAACAAAATCTCTGACAAAAGACAAAGCTTATAGTTACAGGGTGC 114
      |||
      |||
      |||

```


QY 1916 TGGGCTGCATTGTAATACATCTCCCTCTGCATTTTCCCATACATAGAGACTTTGAC 1975
|||||
Db 113 TGGGCTGCATTGTAATACATCTCCCTCTGCATTTTCCCATACATAGAGACTTTGAC 54
|||||
QY 1976 CTGTGAAGCTGCATCTGTATTAAGTAATATCCCAATAA 2016
|||||
Db 53 CTGTGAAGCTGCATCTGTATTAAGTAATATCCCAATAA 13
|||||

RESULT 3
BM129080 571 bp mRNA linear EST 12-MAR-2002
LOCUS 1117C04.Y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
DEFINITION cdna clone IMAGE:5676655 5' similar to TR:09YR3 Q9YR3
N-ACETYLGLUCOSAMINE 6-O-SULFOTRANSFERASE. ; mRNA sequence.

ACCESSION BM129080 GI:17123632
VERSION BM129080.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 511)
AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemisha,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistein,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,
M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
Other ESTs: 1117C04.x1
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownj@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 434.

FEATURES

SOURCE

1. 571

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5676655"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORT1, Site_1: Not 1;
Site_2: Sal I; Starting library constructed using
SuperScript Plasmid Library kit (Life Technologies). cdna
made by oligo-dt priming. Size-selected by column
fractionation: average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an EcoT of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT

109 a 169 c 143 g 150 t

ORIGIN

Query Match 27.1%; Score 551; DB 13; Length 571;
Best Local Similarity 100.0%; Pred. No. 6,1e-162;
Matches 551; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 88 TGAGTCTAGAGAGTGTAAAGCTGTACTTTCACAGCTTCTGGAGAGAGTGTCTTC 147
Db 1 TGAGTCTAGAGAGTGTAAAGCTGTACTTTCACAGCTTCTGGAGAGAGTGTCTTC 60
QY 148 AAGCCCGCTTTCAGAGTCTTCACATTCAGACATGCTACTGCTAAATAAGAGCT 207
Db 61 AAGCCCGCTTTCAGAGTCTTCACATTCAGACATGCTACTGCTAAATAAGAGCT 120
QY 208 CCGTGTCTTGTGTTTCCAGATGGCATCTTGGCTATCTTCCACATGTACAGCA 267
Db 121 CCGTGTCTTGTGTTTCCAGATGGCATCTTGGCTATCTTCCACATGTACAGCA 180
QY 268 CAACATCAGCTCCCTGTCTATGAAGGACACAGCCGAGGATGACAGTGTCTGTCTGTC 327
Db 181 CAACATCAGCTCCCTGTCTATGAAGGACACAGCCGAGGATGACAGTGTCTGTCTGTC 240
QY 328 TTCTCGGCGCTCTGCTCTTCTTTTGTGGGACAGCTTTTGGGACACCCAGATGTTT 387
Db 241 TTCTCGGCGCTCTGCTCTTCTTTTGTGGGACAGCTTTTGGGACACCCAGATGTTT 300
QY 388 CTACCTGATGAGAGCCGCTGAGAGTGTGATGACCTTCAAGCAGAGACCCGCTGGAT 447
Db 301 CTACCTGATGAGAGCCGCTGAGAGTGTGATGACCTTCAAGCAGAGACCCGCTGGAT 360
QY 448 GCTGCACATGCGTGTGGGAGATCTGATACGCGCGCTCTTGTGCGACATGAGCGTCTT 507
Db 361 GCTGCACATGCGTGTGGGAGATCTGATACGCGCGCTCTTGTGCGACATGAGCGTCTT 420
QY 508 TGATGCCATACATGAGACCTGTGTCGCGAGACAGCCGCTTCAAGTGGAGACAG 567
Db 421 TGATGCCATACATGAGACCTGTGTCGCGAGACAGCCGCTTCAAGTGGAGACAG 480
QY 568 CCGGCGCTGTGTTCTGTGACCTGCTGTGACATCATCCACAGATGATATCATCCCCG 627
Db 481 CCGGCGCTGTGTTCTGTGACCTGCTGTGACATCATCCACAGATGATATCATCCCCG 540
QY 628 GGCTTACTGCA 638
Db 541 GGCTTACTGCA 551

RESULT 4

LOCUS

BM128370/c

1113d12.x1

553 bp mRNA linear EST 12-MAR-2002

Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens

DEFINITION

cdna clone IMAGE:5676335 3', mRNA sequence.

ACCESSION

BM128370

BM128370.1

GI:17122922

VERSION

EST.

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS

1 (bases 1 to 553)
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,
Lemisha,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blistein,A.,
Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas,
M., Gibbons,M., McCann,R., Cole,R., Tsagarisvilli,R., Williams,T.,
Jackson,Y. and Bowers,Y.

TITLE

Unpublished (2000)

JOURNAL

Endocrine Pancreas Consortium

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557

Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownj@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 425.

FEATURES

source

Location/Qualifiers

1. 553

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5676553"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORI1; Site:1; Not 1;
Site:2: Sal 1; Starting library constructed using
Superscript Plasmid Library Kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecol of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT 163 a 111 c 127 g 152 t
ORIGIN

Query Match

23.5%; Score 478; DB 13; Length 553;

Best Local Similarity 100.0%; Pred. No. 3.3e-139;

Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1540 TATGTGAGCAGACATCCGACAGATGAAGAGGATTCCTCTTCTTTCTTGTATCT 1599
DB 553 TATGTGAGCAGACATCCGACAGATGAAGAGGATTCCTCTTCTTTCTTGTATCT 494
QY 1600 TCTGTCTGGGACAGATTCAGAGACTTGTGCGCTGGAGGCGCTATTAAACAGACACAGT 1659
DB 493 TCTGTCTGGGACAGATTCAGAGACTTGTGCGCTGGAGGCGCTATTAAACAGACACAGT 434
QY 1660 ATCAGTGAATGATGATTAACCTCCCTGTCCACATCTTCCCAATGGGGAATGATCT 1719
DB 433 ATCAGTGAATGATGATTAACCTCCCTGTCCACATCTTCCCAATGGGGAATGATCT 374
QY 1720 TTCACCAAGAGCTCACAGATTTTCCACAGATGCAATTCCTGAGCCCTTGGAGTTC 1779
DB 373 TTCACCAAGAGCTCACAGATTTTCCACAGATGCAATTCCTGAGCCCTTGGAGTTC 314
QY 1780 CCACTGATTCAG 1839
DB 313 CCACTGATTCAG 254
QY 1840 TCACAGTATCGGTAATCAAAATATGAACAAATATCTGCAACAAAGAGAGAGAGCTT 1899
DB 253 TCACAGTATCGGTAATCAAAATATGAACAAATATCTGCAACAAAGAGAGAGAGCTT 194
QY 1900 AAGTTCACAGAGTGCCTGGGCTGCATTTGAATATCACTTCCCTGCAATTTCCATCA 1959
DB 193 AAGTTCACAGAGTGCCTGGGCTGCATTTGAATATCACTTCCCTGCAATTTCCATCA 134
QY 1960 CATGGAAGACTTTGACTGTGAAAGCTGCCATCTGTATATCTAAATTTCCCAATAG 2017
DB 133 CATGGAAGACTTTGACTGTGAAAGCTGCCATCTGTATATCTAAATTTCCCAATAG 76

RESULT 5
BM128831/c

LOCUS BM128831 553 bp mRNA linear EST 12-MAR-2002
DEFINITION 117c04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
ACCESSION CDNA clone IMAGE:5676553, mRNA sequence.
VERSION BM128831
KEYWORDS BM128831.1 GI:17123383
SOURCE EST.
ORGANISM human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 553)
REFERENCE Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Keestner,K.,
Lemishka,I., Scaerce,M., Brestelli,J., Gradwohl,G., Clifton,S.,
Hillier,L., Maria,M., Pape,D., Wylie,T., Martin,J., Blistain,A.,
Schmitt,A., Phelising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas
M., Gibbons,M., McCann,R., Cole,R., Tsagarishvili,R., Williams,T.,
Jackson,Y. and Bowers,Y.
Endocrine Pancreas Consortium
Unpublished (2000)
CONTACT: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8537
Email: dmelton@biohp.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownj@fas.harvard.edu) This sequence now available from the IMAGE
consortium, for clone orders contact: info@image.llnl.gov
High quality sequence stop: 425.

Location/Qualifiers

FEATURES

source

1. 553

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="5676553"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORI1; Site:1; Not 1;
Site:2: Sal 1; Starting library constructed using
Superscript Plasmid Library Kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column
fractionation; average insert size 1.08 kb. Library was
amplified once on solid support and plasmid DNA from
library was prepared. The library DNA was normalized by
method #4 from Bonaldo, Lennon, and Soares 1996 Genome
Research 6:791-806; 0.5 microgram single-stranded library
plasmid DNA was mixed with 5 micrograms PCR product
representing library inserts and hybridized to an Ecol of
20. Single-stranded (unhybridized) plasmids were isolated
by hydroxyapatite chromatography and used to make this
library."

BASE COUNT 163 a 111 c 127 g 152 t
ORIGIN

Query Match

23.5%; Score 478; DB 13; Length 553;

Best Local Similarity 100.0%; Pred. No. 3.3e-139;

Matches 478; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1540 TATGTGAGCAGACATCCGACAGATGAAGAGGATTCCTCTTCTTGTATCT 1599
DB 553 TATGTGAGCAGACATCCGACAGATGAAGAGGATTCCTCTTCTTGTATCT 494
QY 1600 TCTGTCTGGGACAGATTCAGAGACTTGTGCGCTGGAGGCGCTATTAAACAGACACAGT 1659
DB 493 TCTGTCTGGGACAGATTCAGAGACTTGTGCGCTGGAGGCGCTATTAAACAGACACAGT 434
QY 1660 ATCAGTGAATGATGATTAACCTCCCTGTCCACATCTTCCCAATGGGGAATGATCT 1719

```

Db 433 ATCAGTGGAAATGATCCATAAAACCTCCCTGTCACATCTTGCCCAATGGGAATGATCT 374
OY 1720 TTACCAAAAGAGCTCACCAGCATTTTCCACAGAGATGCAAAATTTGAGCCCTTGAGTTC 1779
Db 373 TTACACCAAGAGCTCACCAGCATTTTCCACAGAGATGCAAAATTTGAGCCCTTGAGTTC 314
OY 1780 CCAGTGGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1839
Db 313 CCAGTGGATTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 254
OY 1840 TCACAGCATTCGTTATATCAGAAATATGAAACAAATCTCTGCACAAAGAGAGAGCTCTT 1899
Db 253 TCACAGCATTCGTTATATCAGAAATATGAAACAAATCTCTGCACAAAGAGAGAGCTCTT 194
OY 1900 AAGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1959
Db 193 AAGTTCACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 134
OY 1960 CATAGAGACTTTGACCTGTGACAGCTGCCATCTGTTAATCTAAATTCACAAATAG 2017
Db 133 CATAGAGACTTTGACCTGTGACAGCTGCCATCTGTTAATCTAAATTCACAAATAG 76

RESULT 6
AL709927 568 bp mRNA linear EST 22-MAR-2002
LOCUS DKFZP686O2364.f1.686 (synonym: h1cc3) Homo sapiens cDNA clone
DEFINITION DKFZP686O2364.5, mRNA sequence.
ACCESSION AL709927
VERSION AL709927.1 GI:19693282
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 668)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and Wiemann
,S.
TITLE Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
MIPS
Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
sequenced by Olgien (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZP686O2364) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcententrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..668
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZP686O2364"
/clone_lib="686 (synonym: h1cc3)"
/tissue_type="human skeletal muscle"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pUC19; Site_1: SfiI; Site_2: SfiI;
cDNA-collection"
BASE COUNT 132 a 202 c 178 g 154 t 2 others
ORIGIN
Query Match 23.0%; Score 468; DB 9; Length 668;
Best Local Similarity 100.0%; Pred. No. 3.7e-136;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 160 CAAGGCTTCACCTCAGCACAATGCTACTGCTAAATAATGAGGCTCCTGCTTTCT 219
|||||

```

```

Db 72 CAAGTCTTCCACTTCAGCACAATGCTACTGCTTAAAAAATGAAAGCTCCTGCTTTCT 131
OY 220 GGTTCACAGATGGCATCTTTGGCTATATCTTCACATGATGACGACCAACATCAGCTC 279
Db 132 GGTTCACAGATGGCATCTTTGGCTATATCTTCACATGATGACGACCAACATCAGCTC 191
OY 280 CCTGCTATGAAGGACAGACCCGAGCGGATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 339
Db 192 CCTGCTATGAAGGACAGACCCGAGCGGATGACAGTGGTGGTGGTGGTGGTGGTGGTGGTGG 251
OY 340 TGGCTCTTCTTTTGGGGAGCTTTTGGGACAGCCACAGATGTTTCTACCTGATGA 399
Db 252 TGGCTCTTCTTTTGGGGAGCTTTTGGGACAGCCACAGATGTTTCTACCTGATGA 311
OY 400 GCCCGCTGGCAGCTGTGATGACCTTCAGCAGAGCAGCCCTGGATGCTGCACATGGC 459
Db 312 GCCCGCTGGCAGCTGTGATGACCTTCAGCAGAGCAGCCCTGGATGCTGCACATGGC 371
OY 460 TGTGGGAGATCTGATACAGGCGGCTCTTGTGGAGATGAGCGCTTTGATGGCTACAT 519
Db 372 TGTGGGAGATCTGATACAGGCGGCTCTTGTGGAGATGAGCGCTTTGATGGCTACAT 431
OY 520 GGAACCTGTGCTCCCGAGACAGTCACCTCTTTCAGTGGAGAGACCGGGCCCTGTG 579
Db 432 GGAACCTGTGCTCCCGAGACAGTCACCTCTTTCAGTGGAGAGACCGGGCCCTGTG 491
OY 580 TTGCGACCTGCTGTGATCATCATCCACAGATGAATCTATCCCGG 627
Db 492 TTGCGACCTGCTGTGATCATCATCCACAGATGAATCTATCCCGG 539

RESULT 7
BF878439 680 bp mRNA linear EST 17-JAN-2001
LOCUS BF878439
DEFINITION M80-ET0109-191100-002-h06 ET0109 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF878439
VERSION BF878439.1 GI:12268569
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 680)
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,
Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare
,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel.: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LIRC Human Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=M80&t=M80-ET0109-
191100-002-h06&t3=2000-11-19&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 17
High quality sequence stop: 641.
FEATURES
source
1..680
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="ET0109"

```

```
/dev_stage="Adult"
/clone_lib="NCI-CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: p173D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAT purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT 168 a 164 c 163 g 184 t 1 others
ORIGIN

Query Match 15.6%; Score 318; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 1.8e-89;
Matches 318; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1468 TGTGTCCACAGCTGTCTACAGAGAGACTTTGTGTCTGCTTGTCTAAGAAACAG 1527
DB 106 TGTGTCCACAGCTGTCTACAGAGAGACTTTGTGTCTGCTTGTCTAAGAAACAG 165
QY 1528 ACTGGGAACTTATGTGACACACATCCACAGAGAGAGGATTTGCTCTCTTC 1587
DB 166 ACTGGGAACTTATGTGACACACATCCACAGAGAGAGGATTTGCTCTCTTC 225
QY 1588 TTTTCTTGATCTTCTGCTGTGGGACACTTCAGAGACTTTGTGCTCGAGGCTATTA 1647
DB 226 TTTTCTTGATCTTCTGCTGTGGGACACTTCAGAGACTTTGTGCTCGAGGCTATTA 285
QY 1648 GCACGACAGATTCAGTGGATTTGATTCATTAACCTCCCTGTCACATCTTCCCAATG 1707
DB 286 GCACGACAGATTCAGTGGATTTGATTCATTAACCTCCCTGTCACATCTTCCCAATG 345
QY 1708 GGGAAATGGATCTTTCACCAAGAGAGCTCCAGAGATTTTCCACAGAGATGCAATTTCTGAG 1767
DB 346 GGGAAATGGATCTTTCACCAAGAGAGCTCCAGAGATTTTCCACAGAGATGCAATTTCTGAG 405
QY 1768 CCCCTGGAGTTCCAGTGC 1785
DB 406 CCCCTGGAGTTCCAGTGC 423

RESULT 8
AM002418 362 bp mRNA linear EST 27-OCT-1999
LOCUS W61903.x1 NCI-CGAP_GC6 Homo sapiens CDNA clone IMAGE:2524564 3',
DEFINITION mRNA sequence.
ACCESSION AM002418
VERSION AM002418.1 GI:5849334
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaldo, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 362
/organism="Homo sapiens"
/db_xref="taxon:9606"

```
/clone_lib="NCI-CGAP_GC6"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: p173D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; Plasmid DNA
from the normalized library NCI CGAP GC4 was prepared, and
ss circles were made in vitro. Following HAT purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clonoids
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."
```

BASE COUNT 108 a 69 c 77 g 108 t
ORIGIN

Query Match 14.2%; Score 288; DB 10; Length 362;
Best Local Similarity 100.0%; Pred. No. 6.3e-80;
Matches 288; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1730 AGCTCACAGCAATTTCCACAGAGATGCAATTCGAGCCCTTGGAGTCCAGTGAAT 1789
DB 362 AGCTCACAGCAATTTCCACAGAGATGCAATTCGAGCCCTTGGAGTCCAGTGAAT 303
QY 1790 CAAGGAAGAGAGTGGGAACAGAGTTGATGCTTACTTATGAGCTTGACATCAGACTAT 1849
DB 302 CAAGGAAGAGAGTGGGAACAGAGTTGATGCTTACTTATGAGCTTGACATCAGACTAT 243
QY 1850 CGGTATCGAATATGAAACAAATCTCTGACAAAGAGAGCAAGCTTAACTTCAAG 1909
DB 242 CGGTATCGAATATGAAACAAATCTCTGACAAAGAGAGCAAGCTTAACTTCAAG 183
QY 1910 GGTGCTGGGCTGCAATTTGAAATATCACTCCCTCGATTTTCCATCAGATAGAGAC 1969
DB 182 GGTGCTGGGCTGCAATTTGAAATATCACTCCCTCGATTTTCCATCAGATAGAGAC 123
QY 1970 TTTGACCTGTGAAGCTGCCATCTGTTAATACATAAATTTCCCAATTAAG 2017
DB 122 TTTGACCTGTGAAGCTGCCATCTGTTAATACATAAATTTCCCAATTAAG 75

RESULT 9
AM572390 419 bp mRNA linear EST 13-MAR-2000
LOCUS XU09A06.x2 NCI-CGAP_Col4 Homo sapiens CDNA clone IMAGE:2795634 3',
DEFINITION mRNA sequence.
ACCESSION AM572390
VERSION AM572390.1 GI:7237123
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
TITLE Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 343.
Location/Qualifiers
1. 419
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="IMAGE:2799634"
/clone.lib="NCI_CGAP_G014"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site:1; Salt:
Site:2: NCI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 Kb. Life Technologies catalog #:
11531-019"

BASE COUNT      122 a      78 c      89 g      130 t
ORIGIN
Query Match      13 98% Score 283; DB 10; Length 419;
Best Local Similarity 99.7%; Pred. No. 2e-78;
Matches 333; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1684 TCCCTGTCACATCTTGGCCCAATGGGGAATGGATGATCTTCCACCAAGAGCTCCACGACATT 1743
DB 419 TCCCTGTCACATCTTGGCCCAATGGGGAATGGATGATCTTCCACCAAGAGCTCCACGACATT 360
QY 1744 TTCCACAGAGATGCAAAATTTCTAGCCCTTGGAGTTCCCAAGTGGATTCAAGAGAGAGAGT 1803
DB 359 TTCCACAGAGATGCAAAATTTCTAGCCCTTGGAGTTCCCAAGTGGATTCAAGAGAGAGAGT 300
QY 1804 GGAACAAGTGGATGCTTACTTATGACCTTATGACCATGACATGACATGCTATGCTATGCAAAAT 1863
DB 299 GGAACAAGTGGATGCTTACTTATGACCTTATGACCATGACATGACATGCTATGCTATGCAAAAT 240
QY 1864 ATGAACAACAATCTCTGCACAAAGAGCAAGCTCTTAAGTTCACAGGGTGGCTGGCTGC 1923
DB 239 ATGAACAACAATCTCTGCACAAAGAGCAAGCTCTTAAGTTCACAGGGTGGCTGGCTGC 180
QY 1924 ATTGAATATGCTCTCCCTCGCATTTTCCATGACATGACATGAGAGCTTGGCTGGAAG 1983
DB 179 ATTGAATATGCTCTCCCTCGCATTTTCCATGACATGACATGAGAGCTTGGCTGGAAG 120
QY 1984 CTGCCATCTGTTAATACTAATAATTCACAAATAG 2017
DB 119 CTGCCATCTGTTAATACTAATAATTCACAAATAG 86

RESULT 10
AM572510/c 417 bp mRNA linear EST 13-MAR-2000
LOCUS      xq18g11.x2 NCI_CGAP_Ut1 Homo sapiens cDNA clone IMAGE:2751044 3'
DEFINITION mRNA sequence.
ACCESSION  AM572510
VERSION     AM572510
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 417)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/HLNtL at:
image.llnl.gov/image/html/lifresources.shtml
Seq primer: -40UP from Gibco
High quality sequence stop: 416.
Location/Qualifiers
1. 417
/organism="Homo sapiens"

FEATURES
Source

/db_xref="taxon:9606"
/clone="IMAGE:2751044"
/clone.lib="NCI_CGAP_Ut1"
/tissue_type="well-differentiated endometrial
adenocarcinoma, 7 pooled tumors"
/lab_host="DH10B"
/notes="Organ: uterus; Vector: pCMV-SPORT6; Site:1; Salt:
Site:2: NCI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.75 Kb. Life Technologies catalog #:
11538-014"

BASE COUNT      140 a      65 c      80 g      132 t
ORIGIN
Query Match      7.4% Score 150; DB 10; Length 417;
Best Local Similarity 100.0%; Pred. No. 5.1e-37;
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1868 AACAAATCTCTGCACAAAGAGCAAGCTCTTAAGTTCACAGAGTGGCTGGCTGCAATT 1927
DB 383 AACAAATCTCTGCACAAAGAGCAAGCTCTTAAGTTCACAGAGTGGCTGGCTGCAATT 324
QY 1928 GAATATCACTTCCCTCGCATTTTCCATGACATGACATGAGCTTGAAGCTGAGAGCTGC 1987
DB 323 GAATATCACTTCCCTCGCATTTTCCATGACATGACATGAGCTTGAAGCTGAGAGCTGC 264
QY 1988 CATCTGTTAATACTAATAATTCACAAATAG 2017
DB 263 CATCTGTTAATACTAATAATTCACAAATAG 234

RESULT 11
BF056840/c 358 bp mRNA linear EST 16-OCT-2000
LOCUS      7k10d12.x1 NCI_CGAP_G06 Homo sapiens cDNA clone IMAGE:3443783 3'
DEFINITION mRNA sequence.
ACCESSION  BF056840
VERSION     BF056840
KEYWORDS    EST.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 358)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLES     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL     Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
Email: cgaps-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldi, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/HLNtL, send email to:
info@image.llnl.gov
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. 358
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3443783"
/clone.lib="NCI_CGAP_G06"
/tissue_type="pooled germ cell tumors"
/lab_host="DH10B"
/notes="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site:1: Not I; Site:2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP_G04 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool

```

of 5,000 clones made from the same library (clonoids 1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT

107 a 68 c 62 g 101 t

Query Match

6.7%; Score 136; DB 12; Length 358;

Best Local Similarity 100.0%; Pred. No. 1.3e-32; Mismatches 0; Indels 0; Gaps 0;

Matches 136; Conservative 0;

QY

1730 AGCTCACACACATTTTCCACAGAGAGCAATTTCTGAGCCCTTGAGCTTCCACGTGATT 1789

DB

358 AGCTCACACACATTTTCCACAGAGAGCAATTTCTGAGCCCTTGAGCTTCCACGTGATT 299

QY

1790 CAAGGAAGAGAGTGGACAAGGTTGGAGTCTTATGAGCTTGACATCACAGCTAT 1849

DB

298 CAGGAAGAGAGTGGACAAGGTTGGAGTCTTATGAGCTTGACATCACAGCTAT 239

QY

1850 CGGTATCAGAAATAT 1865

DB

238 CGGTATCAGAAATAT 223

RESULT 12

BF197521/c

525 bp mRNA linear EST 03-NOV-2000

LOCUS

7084408.x1 NCI-CGAP_Kid11 Homo sapiens cDNA clone IMAGE:3642903 3'

DEFINITION

similar to TR:09UBY3 Q9UBY3 N-ACETYLGLUCOSAMINE

ACCESSION

BF197521

VERSION

BF197521.1 GI:11086670

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS

1 (bases 1 to 525)

TITLE

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: infoimage.llnl.gov
High quality sequence stop: 451.

FEATURES

Location/Qualifiers

Source

1..525

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:3642903"

/lab_host="NCI CGAP_Kid11"

/note="Organ: kidney; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site: Not I; Site2: Eco RI; Plasmid DNA from the normalized library NCI-CGAP_Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500352-1502833). Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

90 a 163 c 187 g 85 t

ORIGIN

Query Match

2.9%; Score 59; DB 12; Length 525;

Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

678 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 736

DB

314 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 256

Best Local Similarity 100.0%; Pred. No. 9.1e-09; Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

678 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 736

DB

260 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 202

RESULT 13

AT824100/c

LOCUS

AT824100

DEFINITION

wj46c01.x1 NCI-CGAP_Lu19 Homo sapiens cDNA clone IMAGE:2405856 3' similar to TR:075667.075667.D071116.4; contains PIR5.b2 PIR5 repetitive element; mRNA sequence.

ACCESSION

AT824100

VERSION

AT824100.1 GI:5444771

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

AUTHORS

1 (bases 1 to 620)

TITLE

NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

JOURNAL

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

COMMENT

Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/dbp/image/image.html
Insert length: 917 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence stop: 490.

FEATURES

Location/Qualifiers

Source

1..620

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:2405856"

/lab_host="NCI CGAP_Lu19"

/tissue="squamous cell carcinoma, poorly differentiated (4 pooled tumors, including primary and metastatic)"

/dev-stage="adult"

/note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; 1st strand cDNA was prepared from pooled lung tumor tissue, and was then primed with a Not I - oligo(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT

106 a 195 c 227 g 89 t 3 others

ORIGIN

Query Match

2.9%; Score 59; DB 9; Length 620;

Best Local Similarity 100.0%; Pred. No. 8e-09; Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

678 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 736

DB

314 AGGCTGGCGGCTTACAGCCAGCGTGGCTTCAAGAGGTGGCTTTCACCTGCAG 256

RESULT 14

BE857538/c

found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.lnl.gov/dbp/image.html

Seq primer: -400p from gldco
High quality sequence stop: 431.

FEATURES

Location/Qualifiers
1. 436

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI_CGAP_Co20"
/issue_type="moderately differentiated adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: PCMV-SPORT6; Site:1: SalI;
Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Normalized to Cot 500. Average insert size 1.11kb
Normalized version of NCI_CGAP_Co18. Library constructed
by Life Technologies."

BASE COUNT

64 a 136 c 169 g 67 t

ORIGIN

Query Match 2.3%; Score 47; DB 10; Length 436;

Best Local Similarity 100.0%; Pred. No. 5.7e-05;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 678 AGGCGTCCGCTCTACACGACGCTGCTCAAGAGTGCGCTTC 724
|||||
DB 309 AGGCGTCCGCTCTACACGACGCTGCTCAAGAGTGCGCTTC 263

RESULT 17

BM246681 497 bp mRNA linear EST 01-FEB-2002
K0741E04-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
LOCUS BM246681.1 GI:17881951
DEFINITION CDNA Library (Long) Mus musculus CDNA clone K0741E04 3', mRNA
sequence.

ACCESSION BM246681
VERSION BM246681.1 GI:17881951
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 497)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Alba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)

JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0741 row: E column: 04
Seq primer: -21M13 forward
High quality sequence stop: 497
POLYA=Yes

FEATURES

Location/Qualifiers
1. 497

/organism="Mus musculus"
/strain="C57BL/6NCR"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)"
/issue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx. 10 weeks old"
/lab_host="DH10B"
/note="Vector: pSPORT1 (Invitrogen); Site:1: SalI; Site:2:
NotI; Mouse CDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research

Program, NIH (<http://igsun.grc.nia.nih.gov/cdna>). This is
a long-transcript enriched CDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544191]) Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National Cancer
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer [Invitrogen:
5'-GGAGTGTCTGATGACGACGCGCCGCTTTTCTTTT-3'] from
0.9 ug of total RNA, treated with R4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to linc-linker L-Sal4, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Tag polymerase (Takara) with a primer Sal4-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."

BASE COUNT

87 a 137 c 134 g 139 t

ORIGIN

Query Match 1.3%; Score 27; DB 13; Length 497;

Best Local Similarity 100.0%; Pred. No. 87;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 494 GACATGAGCGCTCTTGATGCTACATG 520
|||||
DB 441 GACATGAGCGCTCTTGATGCTACATG 467

RESULT 18

BM245312 536 bp mRNA linear EST 01-FEB-2002
K0722H07-3 NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)
LOCUS BM245312.1 GI:17880582
DEFINITION CDNA Library (Long) Mus musculus CDNA clone K0722H07 3', mRNA
sequence.

ACCESSION BM245312
VERSION BM245312.1 GI:17880582
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 536)
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Alba, K., Taub, D., Longo, D.L., Keller, J., and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)

JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0722 row: H column: 07
Seq primer: -21M13 forward
High quality sequence stop: 536
POLYA=Yes

FEATURES

Location/Qualifiers
1. 536

/organism="Mus musculus"
/strain="C57BL/6NCR"
/db_xref="taxon:10090"
/db_xref="taxon:10090"
/clone_lib="NIA Mouse Hematopoietic Stem Cell
(Lin-/c-Kit-/Sca-1-) CDNA Library (Long)"
/issue_type="Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx. 10 weeks old"

/lab_host="DH10B"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (http://19sun.grc.nia.nih.gov/cDNA). This is
a long-transcript enriched cDNA library (Ref. genome Res.
11: 1553-1558 (2001). [PMID: 11544199]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Kallier (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an oligo(dT) primer [Invitrogen:
5'-pGACTAGTCCTGATGCGAGCGCCGCTTTTCTTTTCTT-3'] from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker IL-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Taq polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Plao (NIA)."

BASE COUNT 93 a 150 c 147 g 146 t
ORIGIN

Query Match 1.3%; Score 27; DB 13; Length 536;
Best Local Similarity 100.0%; Pred. No. 82;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCCTACATG 520
|||||
Db 441 GACATGAGCGCTTTGATGCCTACATG 467
|||||

RESULT 19
BG9664671 695 bp mRNA linear EST 12-JUN-2001
LOCUS 602831875F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4986258 5',
DEFINITION mRNA sequence.
ACCESSION BG9664671
VERSION BG9664671.1 GI:14352308
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 695)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: LHAM10995 row: f column: 19
High quality sequence stop: 684.
Location/Qualifiers
1..695
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4986258"
/clone_1lb="NCI_CGAP_Co24"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 118 a 216 c 195 g 166 t
ORIGIN

Query Match 1.3%; Score 27; DB 13; Length 695;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCCTACATG 520
|||||
Db 345 GACATGAGCGCTTTGATGCCTACATG 371
|||||

RESULT 20
BG963298 783 bp mRNA linear EST 12-JUN-2001
LOCUS 60282716F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4982643 5',
DEFINITION mRNA sequence.
ACCESSION BG963298
VERSION BG963298.1 GI:14350935
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathu; Muridae; Murinae; Mus.
1 (bases 1 to 783)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL at:
http://image.llnl.gov
Plate: LHAM10985 row: p column: 04
High quality sequence stop: 741.
Location/Qualifiers
1..783
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4982643"
/clone_1lb="NCI_CGAP_Co24"
/lab_host="DH10B (TI phage-resistant)"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.6 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 141 a 236 c 226 g 180 t
ORIGIN

Query Match 1.3%; Score 27; DB 13; Length 783;
Best Local Similarity 100.0%; Pred. No. 62;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTTGATGCCTACATG 520
|||||
Db 376 GACATGAGCGCTTTGATGCCTACATG 402
|||||

RESULT 21
BG966340 852 bp mRNA linear EST 12-JUN-2001
LOCUS 602832826F1 NCI_CGAP_Co24 Mus musculus cDNA clone IMAGE:4987327 5',
DEFINITION mRNA sequence.
ACCESSION BG966340
VERSION BG966340.1 GI:14353977

GACGAGAGATCTCGAGTATATTAATTAATCCCCCCCCCCCCC 3']. cDNA was cleaved with XhoI and SstI. Cloning sites, 5' end: XhoI; 3' end: SstI.
Host: SOLR.

FEATURES

source
Location/Qualifiers
1..1923
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:2310003618"
/db_xref="MGI:1904558"
/db_xref="taxon:10090"
/clone="2310003618"
/sex="male"
/tissue.type="tongue"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1923
/gene="Cstf4"
/gene="Cstf4"
96..1262
/note="carbohydrate (chondroitin 6/keratan)
sulfotransferase 4
data source:MGI, source key:MGI:1349479, evidence:ISS
putative"
/codon_start=1
/protein_id="BAB26078.1"
/db_xref="GI:12843702"
/db_xref="MGI:1349479"
/translation="MMLIKRRLMFLGQVYVVALTIHNSVHRHLSORESRPPVY
LVLSMRSGSSFGQLFGQHPDVYLVPEAHVWMTTSTAMKLNHNAVRLNLSVL
CDMSVFAYANNGPRKOSLFEQMSALCAPCDEFPAREHISPPHCKILCCOOP
DMVERKASGCVVYLVKEVRLSLQALPLDPLNLHVVLVADPRAVFRREHTTI
ELMVDSHVILGQLEETKEEDQPYAMKTIICKSOVDYKAIOTLPEALOOYLFLRE
DLVAPLAOTRTLKFGVGLDFPLHLOTMYVYVTEGKGGMGOGAHTNRLNLSQARR
WSLPERKSQLDQSGEMDLGLGYVRSQDEGNSLIDLSSSHILGOVFRFG"

BASE COUNT 445 a 533 c 492 g 453 t
ORIGIN

Query Match 1.3%; Score 27; DB 11; Length 1923;
Best Local Similarity 100.0%; Pred. No. 31;

Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 494 GACATGAGCGCTTGATGCTCATG 520
|||||
Db 405 GACATGAGCGCTTGATGCTCATG 431

RESULT 23
BM036473/c 501 bp mRNA linear EST 05-NOV-2001
LOCUS F01107.X1 Gong zebrafish testis Danio rerio cDNA clone 5377309 3',
DEFINITION mRNA sequence.
ACCESSION BM036473
VERSION BM036473.1 GI:16750044
KEYWORDS EST.
SOURCE zebrafish
ORGANISM Danio rerio

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes
; Cyprinidae; Danio.
1 (bases 1 to 501)
Clark, M., Johnson, S.L., Lebrach, H., Lee, R., Li, F., Maira, M., Eddy,
'S., Hillier, L., Kucaba, T., Martin, J., Beck, C., Mylly, T., Underwood
'K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B.,
Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurk, R., Rilter, E.,
Kohn, S., Shih, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R.,
and Wilson, R.

TITLE Washu zebrafish EST Project 1998
JOURNAL Unpublished (1998)
COMMENT Contact: Stephen L. Johnson
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

FEATURES

Email: zbrafish@watson.wustl.edu
The library was constructed by Dr. Z. Gong. DNA Sequencing by:
Washington University Genome Sequencing Center St. Louis. Please
contact Zhiyuan Gong for further information on this library
(National University of Singapore, Department of Biological
Sciences, Lower Kent Ridge Road, Singapore 119260).
High quality sequence stop: 367.
Location/Qualifiers
1..501
/organism="Danio rerio"
/db_xref="taxon:7955"
/clone="5377309"
/clone_lib="Gong zebrafish testis"
/sex="male"
/dev_stage="4-5 month"
/lab_host="DH10B (phage-resistant)"
/note="Organ: testis (pooled); Vector: pBluescript SK-;
Site_1: XhoI; Site_2: EcoRI; Poly A+ RNA was isolated from
the testes of 31 male adult zebrafish (4-5 month old).
cDNAs were made using oligo-dt primers and inserted into
lambda ZAP II vector (Stratagene) by Dr. Z. Gong, in vivo
mass-excision to pBluescript SK- following the Washington
University protocol
(http://genome.wustl.edu/est/lambda_protocol.shtml).
Please contact Zhiyuan Gong for further information on
this library (National University of Singapore,
Department of Biological Sciences, Lower Kent Ridge Road,
Singapore 119260)."

BASE COUNT 171 a 94 c 90 g 146 t
ORIGIN

Query Match 1.2%; Score 25; DB 13; Length 501;
Best Local Similarity 100.0%; Pred. No. 3; 6e+02;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 CCCAATATAGAAAAA 2032
|||||
Db 29 CCCAATATAGAAAAA 5

RESULT 24
CNS02BK0/c 682 bp DNA linear GSS 12-MAY-2000
LOCUS Tetradon nigroviridis genome survey sequence PUC-01 end of clone
DEFINITION 253N23 of library G from Tetradon nigroviridis, genomic survey
sequence.
ACCESSION AL189971
VERSION AL189971.1 GI:7828075
KEYWORDS GSS; genome survey sequence.
SOURCE Tetradon nigroviridis.
ORGANISM Tetradon nigroviridis

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
Acanthomorphna; Acanthopterygii; Percomorphna; Tetraodontiformes;
Tetraodontidae; Tetraodon.
1 (bases 1 to 682)
Roest-Crolius, H., Jalllon, O., Dasilva, C., Bouneau, L., Fisher, C.,
Barnot, A., Fitzmes, C., Wincker, P., Brottier, P., Quetier, F.,
Saurin, W. and Weissenbach, J.

TITLE Human gene number estimate provided by genome wide analysis using
Tetradon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 682)
Roest-Crolius, H., Jalllon, O., Dasilva, C., Fitzmes, C., Fisher, C.,
Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and
Weissenbach, J.

TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetradon nigroviridis
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 682)
GENOSCOPE.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000)

COMMENT

This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES

source

Location/Qualifiers

1..682

/organism="Tetraodon nigroviridis"

/db_xref="taxon:99883"

/clone="253N23"

/clone.lib="G"

/note="Genoscope sequence ID : C0AG253CG12SP1-end ; PUC-ori"

BASE COUNT 202 a 122 c 113 g 224 t 21 others

ORIGIN

Query Match 1.2%: Score 25; DB 17; Length 682;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2008 CCCCAATAGAGAAAAA 2032

Db 241 CCCCAATAGAGAAAAA 217

Search completed: February 22, 2003, 22:52:05
Job time : 2651 secs

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

(without alignments)
11281.362 Million cell updates/sec

Sequence: 1 ggctcgagccagtgct.....atagaaaaaa 2032

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 21

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Listing first 1000 summaries

Database :

1:	gb_da:*
2:	gb_hlg:*
3:	gb_in:*
4:	gb_om:*
5:	gb_ov:*
6:	gb_pat:*
7:	gb_ph:*
8:	gb_pl:*
9:	gb_pr:*
10:	gb_ro:*
11:	gb_st:*
12:	gb_sy:*
13:	gd_un:*
14:	gd_vl:*
15:	em_ba:*
16:	em_fun:*
17:	em_hum:*
18:	em_in:*
19:	em_mu:*
20:	em_om:*
21:	em_or:*
22:	em_ov:*
23:	em_pat:*
24:	em_ph:*
25:	em_pl:*
26:	em_ro:*
27:	em_sts:*
28:	em_un:*
29:	em_vl:*
30:	em_hlg_hum:*
31:	em_hlg_in:*
32:	em_hlg_om:*
33:	em_hlg_ov:*
34:	em_hlg_pat:*
35:	em_hlg_ph:*
36:	em_hlg_pl:*
37:	em_hlg_pr:*
38:	em_hlg_ro:*
39:	em_hlg_sy:*
40:	em_hlg_un:*
41:	em_hlg_om_*:

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB	ID	Description
1	2032	100.0	2032	6	AR203335	AR203335 Sequence
2	2032	100.0	2032	9	AF131235	AF131235 Homo sapi
3	1806	91.3	2011	9	AK026635	AK026635 Homo sapi
4	1806	88.9	183289	9	AC010547	AC010547 Homo sapi
5	1756	86.4	1992	9	AF280088	AF280088 Homo sapi
6	1211	59.6	1333	9	AF149783	AF149783 Homo sapi
7	505	24.9	517	6	AX381256	AX381256 Sequence
8	286	14.1	389	6	AX381326	AX381326 Sequence
9	59	2.9	1462	9	AF176839	AF176839 Homo sapi
10	59	2.9	1447	6	AX327330	AX327330 Sequence
11	59	2.9	2170	9	AF176838	AF176838 Homo sapi
12	59	2.9	2544	9	AF219990	AF219990 Homo sapi
13	59	2.9	3378	9	AF246718	AF246718 Homo sapi
14	59	2.9	3786	9	AF280086	AF280086 Homo sapi
15	59	2.9	71503	9	AF219991	AF219991 Homo sapi
16	59	2.9	157737	9	AC009163	AC009163 Homo sapi
17	59	2.9	194832	9	AC025287	AC025287 Homo sapi
18	59	2.9	208185	2	AC009105	AC009105 Homo sapi
19	37	1.8	37	6	AR203340	AR203340 Sequence
20	27	1.3	1926	6	AF131236	AF131236 Mus muscu
21	27	1.3	2201	10	AF109155	AF109155 Mus muscu

ALIGNMENTS

LOCUS	AR203335	2032 bp	DNA	linear	PAT 20-JUN-2002
DEFINITION	Sequence 2 from patent US 6365365.				
ACCESSION	AR203335				
VERSION	AR203335.1	GI:21499698			
KEYWORDS					
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	Unclassified.				
AUTHORS	1 (bases 1 to 2032)				
TITLE	Bistrup, A., Rosen, S.D., Tangemann, K. and Hemmerich, S.				
JOURNAL	Method of determining whether an agent modulates glycosyl				
FEATURES	Sulfotransferase-3				
source	Patent: US 6365365-A 2 02-APR-2002;				
	location/Qualifiers				
	1..2032				
BASE COUNT	468 a 569 c 490 g 505 t				
ORIGIN	/organism="unknown"				
Query Match	100.0%; Score 2032; DB 6; Length 2032;				
Best Local Similarity	100.0%; Pred. No. 0;				
Matches 2032; Conservative	0; Mismatches	0; Indels	0; Gaps	0;	
QY	1 GGGCTGAGGCCAGATGCTCCAGTCTGAGGGGAAAAATCTTCCTCATTTGCTTCCACG 60				
Db	1 GGGCTGAGGCCAGATGCTCCAGTCTGAGGGGAAAAATGCTTCCTCATTTGCTTCCACG 60				
QY	61 CCCACTCAACAGTCTCCACCCGCTGAGTCAAGAGTGAAGCGTTCCTTCA 120				
Db	61 CCCAGCTCAACAGTCTCCACCCGCTGAGTCAAGAGTGAAGCGTTCCTTCA 120				
QY	121 CAGCTTCTTGAGAGCGAGTCTTCTCAAGCCGCTGTGCAAGGCTTCCACTTCAGCAC 180				
Db	121 CAGCTTCTTGAGAGCGAGTCTTCTCAAGCCGCTGTGCAAGGCTTCCACTTCAGCAC 180				
QY	181 AATGCTACTGCTAAAAAAATGAAGCTCTCTGTTTGTGTTTCCAGATGCGCACTT 240				
Db	181 AATGCTACTGCTAAAAAAATGAAGCTCTCTGTTTGTGTTTCCAGATGCGCACTT 240				

Db 181 AATGCTACTGCGCTAAATAAATGAAGCTCGTGGCTTTGTTTCCAGATGAGCATTT 240
 QY 241 GGGCTATTCTTCCATGATGACAGCCACAACATGACCTCCCTGTCTATATGAAGCAGCC 300
 Db 241 GGGCTATTCTTCCATGATGACAGCCACAACATGACCTCCCTGTCTATATGAAGCAGCC 300
 QY 301 CGAGCGATGACAGTGTGCTGTGCTTCTGCTGCGGCTGTGCTGCTTCTTGTGGGGA 360
 Db 301 CGAGCGATGACAGTGTGCTGTGCTTCTGCTGCGGCTGTGCTGCTTCTTGTGGGGA 360
 QY 361 GCTTTTGGGAGACCCAGATGTTTCTACCTGATGAGACCCGCTGTGCTGCTGCTGAT 420
 Db 361 GCTTTTGGGAGACCCAGATGTTTCTACCTGATGAGACCCGCTGTGCTGCTGCTGAT 420
 QY 421 GACCTTCAAGCAGAGACCCGCTGATGCTGACATGAGCTGCTGCTGCTGCTGATGAG 480
 Db 421 GACCTTCAAGCAGAGACCCGCTGATGCTGACATGAGCTGCTGCTGCTGCTGATGAG 480
 QY 481 GCTTCTTGTGCGAGATGAGCTGTTTGTATGCTTACATGAGAACTGTGCTCCGAGACA 540
 Db 481 GCTTCTTGTGCGAGATGAGCTGTTTGTATGCTTACATGAGAACTGTGCTCCGAGACA 540
 QY 541 GTCAGGCTTTCAGTGGGAGAAAGCCGCGCTGTGCTGCTGCTGCTGCTGCTGAT 600
 Db 541 GTCAGGCTTTCAGTGGGAGAAAGCCGCGCTGTGCTGCTGCTGCTGCTGCTGAT 600
 QY 541 GTCAGGCTTTCAGTGGGAGAAAGCCGCGCTGTGCTGCTGCTGCTGCTGCTGAT 600
 Db 541 GTCAGGCTTTCAGTGGGAGAAAGCCGCGCTGTGCTGCTGCTGCTGCTGCTGAT 600
 QY 601 CATCCCAAGATGAATCATCCCGCGGCTCATGACAGGCTGCTGCTGCTGCTGCTGAT 660
 Db 601 CATCCCAAGATGAATCATCCCGCGGCTCATGACAGGCTGCTGCTGCTGCTGCTGAT 660
 QY 661 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Db 661 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 QY 721 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 Db 721 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
 QY 781 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 Db 781 CTTTGGAGTGTGAGAGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840
 QY 841 TCTCATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 900
 Db 841 TCTCATGATGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 900
 QY 901 CCAACCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 Db 901 CCAACCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 961 CCAACCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 Db 961 CCAACCTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1020
 QY 1021 TCGAGCCCTGTGAGCCAGATTCCTCCAGATGATGATGATGATGATGATGATGAT 1080
 Db 1021 TCGAGCCCTGTGAGCCAGATTCCTCCAGATGATGATGATGATGATGATGATGAT 1080
 QY 1081 CCAATCTTCAAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 Db 1081 CCAATCTTCAAGACCTGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1140
 QY 1141 CCAACAAATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 Db 1141 CCAACAAATGAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
 QY 1201 TGAAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 Db 1201 TGAAGAGTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1260
 QY 1261 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320
 Db 1261 CCAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1320

QY 1321 TGTCCCTGAGCAATCCACTAAGAGGTTGAGAAAGCTTTTCTGCTCCACCTGTGTGAGCC 1380
 Db 1321 TGTCCCTGAGCAATCCACTAAGAGGTTGAGAAAGCTTTTCTGCTCCACCTGTGTGAGCC 1380
 QY 1381 TCACTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 Db 1381 TCACTGATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
 QY 1441 TGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 Db 1441 TGTGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
 QY 1501 GTGTCCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 Db 1501 GTGTCCATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1560
 QY 1561 CAGTGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Db 1561 CAGTGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 QY 1621 AGACTTGTGAGCTGAGAGGCTTATTAGCAGACAGACAGATGATGATGATGATGAT 1680
 Db 1621 AGACTTGTGAGCTGAGAGGCTTATTAGCAGACAGACAGATGATGATGATGATGAT 1680
 QY 1681 ACCTGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 Db 1681 ACCTGCTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1740
 QY 1741 ATTTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 Db 1741 ATTTCACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1800
 QY 1801 GTGGAGACAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 Db 1801 GTGGAGACAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1860
 QY 1861 AATATGAACAAATCTGCTGACAAAGAGAGAGCTTTAAGTTACAGAGGTGCTGAGC 1920
 Db 1861 AATATGAACAAATCTGCTGACAAAGAGAGAGCTTTAAGTTACAGAGGTGCTGAGC 1920
 QY 1921 TGCATTTGAATATCACTTCCCTGCTGATTTTCCATGATGATGATGATGATGATGAT 1980
 Db 1921 TGCATTTGAATATCACTTCCCTGCTGATTTTCCATGATGATGATGATGATGATGAT 1980
 QY 1981 AAGCTGCCATCTGTTAATCTAAATTCCTAAATTCCTAAATTCCTAAATTCCTAAAT 2032
 Db 1981 AAGCTGCCATCTGTTAATCTAAATTCCTAAATTCCTAAATTCCTAAATTCCTAAAT 2032
 RESULT 2
 AF131235 2032 bp mRNA linear PRI 10-AUG-1999
 LOCUS AF131235
 DEFINITION Homo sapiens N-acetylglucosamine 6-O-sulfotransferase mRNA,
 complete cds.
 ACCESSION AF131235
 VERSION AF131235.1 GI:4927113
 KEYWORDS
 SOURCE
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 2032)
 Bistrup, A., Bhakta, S., Lee, J. K., Belov, Y. Y., Gunn, M. D., Zuo, F. R.,
 Huang, C. C., Kannagi, R., Rosen, S. D., and Hemmerich, S.
 Sulfotransferases of two specificities function in the
 reconstitution of high endothelial cell ligands for L-selectin
 J. Cell Biol. 145 (4), 899-910 (1999)
 JOURNAL MEDLINE
 PUBMED 10330415
 REFERENCE 2 (bases 1 to 2032)
 Bistrup, A., Tangemann, K., Bhakta, S., Lee, J. K., Belov, Y. Y.,
 Gunn, M. D., Zuo, F. R., Huang, C. C., Kannagi, R., Rosen, S. D., and

Hemerich, S.
Direct Submission
JOURNAL
Submitted (25-FEB-1999) Respiratory Diseases, Roche Bioscience,
3401 Hillview Avenue, Palo Alto, CA 94304, USA
FEATURES
source
Location/Qualifiers
1..2032
/organism="Homo sapiens"
/db_xref="taxon:9606"
/cell_type="high endothelial"
/tissue_type="tmsil"
182..1342
/function="biosynthesis of endothelial ligands for
L-selection"
/note="HEC-G1CMAC-6-ST"
/codon_start=1
/product="N-acetylglucosamine 6-O-sulfotransferase"
/protein_id="AAD3015.1"
/db_xref="GI:4927114"
/translation="MLPKKKKLLFLVSQMAILALFFHMYSHNISLSMKAOPERMH
VLVLSMRSGSEVGOLEFGQHPDVEYLMPEAMVMTFQKOSTAMLMAMVRLIRAYE
LCDMSVFAYMEPGPBRSSILFQWENSALCASAPACILIPDEILIPRHHRLISQDP
FEVEKACRSYSHVLKEVREPNQSLYPLKPSLNLHLYLDRPRAVRSERK
GDLMIDRIYMGHOKIKREDOPYVMQVTCOSOLETKTQSLPRLDERLYLVAY
EDLRAFAPOQSRMTEFVGLFELPHLQIOWHNITRGKMDHAETVARALANSQAM
RWSLPYEVKSLQACGDANMLLGRVHRSEQRNDLLDLSTWVPEQIH"
BASE COUNT 468 a 569 c 490 g 505 t
ORIGIN
Query Match 100.0%; Score 2032; DB 9; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCTTGAGCGCAGATGCTCCAGTCTGTGGGGAATAATGCTTCATTTGCTTCCAG 60
Db 1 GGCTGAGCGCAGATGCTCCAGTCTGTGGGGAATAATGCTTCATTTGCTTCCAG 60
QY 61 CCACACTAAGACAGTCTCCACCCCTTGAGCTCAGAGTGAAGCTGTACTTCA 120
Db 61 CCACACTAAGACAGTCTCCACCCCTTGAGCTCAGAGTGAAGCTGTACTTCA 120
QY 121 CAGCTTCTGGGAGGAGTGTCTTCTCAAGCCCGCTTCAAGAGCTTCCACTCAGAC 180
Db 121 CAGCTTCTGGGAGGAGTGTCTTCTCAAGCCCGCTTCAAGAGCTTCCACTCAGAC 180
QY 181 AATGCTACTGCTAAAAAATGAAGCTCTGCTTTCTGCTTCCAGATGCGCATCTT 240
Db 181 AATGCTACTGCTAAAAAATGAAGCTCTGCTTTCTGCTTCCAGATGCGCATCTT 240
QY 241 GCGCTATTTCTCCAGATGTAAGCCACACATCAGTCCCTGCTATGAAGGACAGCC 300
Db 241 GCGCTATTTCTCCAGATGTAAGCCACACATCAGTCCCTGCTATGAAGGACAGCC 300
QY 301 CGAGCGCATGACGTGTGTTCTTCTTCTGCGGCTCTGCTGCTCTTTTGTGGGCA 360
Db 301 CGAGCGCATGACGTGTGTTCTTCTTCTGCGGCTCTGCTGCTCTTTTGTGGGCA 360
QY 361 GCTTTTGGGAGACCCAGATGTTTCTACTATGAAGCCCGCTGGGACGTTGAT 420
Db 361 GCTTTTGGGAGACCCAGATGTTTCTACTATGAAGCCCGCTGGGACGTTGAT 420
QY 421 GACCTTCAGAGACACCGCTGATGCTGACATGAGTGGGATGATGATGAGGAC 480
Db 421 GACCTTCAGAGACACCGCTGATGCTGACATGAGTGGGATGATGATGAGGAC 480
QY 481 CGTCTTCTGTGCGACATGAGCTTTGATGCTACATGGAACCTGCTCCCGAGACA 540
Db 481 CGTCTTCTGTGCGACATGAGCTTTGATGCTACATGGAACCTGCTCCCGAGACA 540
QY 541 GTCCAGCCTTTCAAGTGGAGAACAGCCGGCCCTGTGCTCAGCTGCTGTGACAT 600
Db 541 GTCCAGCCTTTCAAGTGGAGAACAGCCGGCCCTGTGCTCAGCTGCTGTGACAT 600
QY 601 CATCCCAAGATGAATCATCCCCCGGCTCACTGACAGGCTCCTGTGACATCAAGCC 660

Db 601 CATCCCAAGATGAATCATCCCCCGGCTCACTGACAGGCTCCTGTGACATCAAGCC 660
QY 661 CTTGAGCTGTGAGAGAGCGCTGCGCTCTACAGCCACGTGTGCTCAAGAGGTCG 720
Db 661 CTTGAGCTGTGAGAGAGCGCTGCGCTCTACAGCCACGTGTGCTCAAGAGGTCG 720
QY 721 CTTCTCACTGACAGTCCCTCTACCCGCTGCTGAAGACCCCTCCCTCAACCTGAT 780
Db 721 CTTCTCACTGACAGTCCCTCTACCCGCTGCTGAAGACCCCTCCCTCAACCTGAT 780
QY 781 CGTGACCTGTGTCGGGAGACCCCGGCGCGTCTCCGTCCTCCGAGAACGACAAAGGAGA 840
Db 781 CGTGACCTGTGTCGGGAGACCCCGGCGCGTCTCCGTCCTCCGAGAACGACAAAGGAGA 840
QY 841 TCTCATGATTGACAGTGCATTTGATGGGACAGCATGACGAAACTCAAGAGAGAGA 900
Db 841 TCTCATGATTGACAGTGCATTTGATGGGACAGCATGACGAAACTCAAGAGAGAGA 900
QY 901 CCAACCCCTACTATGTGATGAGAGGTCATCTGCCAAGCCAGCTGAGATCTTCAAGACCAT 960
Db 901 CCAACCCCTACTATGTGATGAGAGGTCATCTGCCAAGCCAGCTGAGATCTTCAAGACCAT 960
QY 961 CCACTGCTTGCCCAAGGCGCTGCAGAGACGCTACCTGCTTGCGCTATGAGAGACCTGGC 1020
Db 961 CCACTGCTTGCCCAAGGCGCTGCAGAGACGCTACCTGCTTGCGCTATGAGAGACCTGGC 1020
QY 1021 TCGAGCCCTGTGCGCCAGACTTCCGATGATGAATTGCTGGATGGAATTTCTGCC 1080
Db 1021 TCGAGCCCTGTGCGCCAGACTTCCGATGATGAATTGCTGGATGGAATTTCTGCC 1080
QY 1081 CCACTTTCAGACCTGGGAGCATTAACATCACCCGAGGAGCATGGGTGACACCGCTTT 1140
Db 1081 CCACTTTCAGACCTGGGAGCATTAACATCACCCGAGGAGCATGGGTGACACCGCTTT 1140
QY 1141 CCACACAAATGCCAGGAGTGCCTTAAATGCTCCAGAGCTGGCGCTGCTTTGGCCCTA 1200
Db 1141 CCACACAAATGCCAGGAGTGCCTTAAATGCTCCAGAGCTGGCGCTGCTTTGGCCCTA 1200
QY 1201 TGAAGAAGTTCTGCACTTCAAGAAAGCCTGGGAGATGCCATGAATTTGCTGGGCTACCG 1260
Db 1201 TGAAGAAGTTCTGCACTTCAAGAAAGCCTGGGAGATGCCATGAATTTGCTGGGCTACCG 1260
QY 1261 CCACGTGAGATCTGAACAGAGAGAGAAACCTGTTGCTGATCTTGTCTACCTGAC 1320
Db 1261 CCACGTGAGATCTGAACAGAGAGAGAAACCTGTTGCTGATCTTGTCTACCTGAC 1320
QY 1321 TGTCCCTGAGCAAAATCCACTAAGAGGGTTGAGAAAGCTTGTGCTGCCACCTGCTGACGC 1380
Db 1321 TGTCCCTGAGCAAAATCCACTAAGAGGGTTGAGAAAGCTTGTGCTGCCACCTGCTGACGC 1380
QY 1381 TCACTCACTTCTGTAATGCTTGTGAGCCTTGCTCTGATCTGAGCCTTAACATG 1440
Db 1381 TCACTCACTTCTGTAATGCTTGTGAGCCTTGCTCTGATCTGAGCCTTAACATG 1440
QY 1441 TCTGTGGGTATCACACTGAGAGTGTGATGTGTCCACAGCTGCTCAAGCAGAGACTTTT 1500
Db 1441 TCTGTGGGTATCACACTGAGAGTGTGATGTGTGTCCACAGCTGCTCAAGCAGAGACTTTT 1500
QY 1501 GTGTCCATGCTGTGTGTAGAAAAACAGACGAGGGAACCTATGTGACACACATCCAC 1560
Db 1501 GTGTCCATGCTGTGTGTAGAAAAACAGACGAGGGAACCTATGTGACACACATCCAC 1560
QY 1561 CAGTGAAGAGGGATATGCTCTTCTTTCTTGAATCTTCTGCTGAGGACACTTCAAG 1620
Db 1561 CAGTGAAGAGGGATATGCTCTTCTTTCTTGAATCTTCTGCTGAGGACACTTCAAG 1620
QY 1621 AGACTTTGTGGCCCTGAGGAGCTATTAAGCAGACAGAGTACAGTGAATGATCAATA 1680
Db 1621 AGACTTTGTGGCCCTGAGGAGCTATTAAGCAGACAGAGTACAGTGAATGATCAATA 1680
QY 1681 ACCTCCCTGTCCACATCTTCCCAATGGGAGAGATCTTCCAAAGAGCTCACACAG 1740

Db 1681 ACCTCCCTGTCACATCTTGGCCCAATGGGGAATGATCTTTTCACCAAGAAGCTCACACAGC 1740

QY 1741 ATTTTCCACAGAGATCAAAATTTCTGAGCCCTTGAGATCCCAAGTATTCAGGAAGGA 1800

Db 1741 ATTTTCCACAGAGATCAAAATTTCTGAGCCCTTGAGATCCCAAGTATTCAGGAAGGA 1800

QY 1801 GTGGGAACAAGTTGGATGCTCTATTATGAGCTTGACATCAGATTCAGGAATCAGA 1860

Db 1801 GTGGGAACAAGTTGGATGCTCTATTATGAGCTTGACATCAGATTCAGGAATCAGA 1860

QY 1861 AATATGAACAACAAATCTCTGACAAAAAGCAAGCTCTTAAGTTCAAGAGGTGCTGGGC 1920

Db 1861 AATATGAACAACAAATCTCTGACAAAAAGCAAGCTCTTAAGTTCAAGAGGTGCTGGGC 1920

QY 1921 TGCATTGATATCAGCTTCCCTCTGATTTTCCCATCAGATGAGAGCTTTGACCTGTG 1980

Db 1921 TGCATTGATATCAGCTTCCCTCTGATTTTCCCATCAGATGAGAGCTTTGACCTGTG 1980

QY 1981 AAGCTGCATCTGTATTAATACTAAATTTCCCAATAAGAAAAA 2032

Db 1981 AAGCTGCATCTGTATTAATACTAAATTTCCCAATAAGAAAAA 2032

RESULT 3

AK026635

LOCUS

DEFINITION

AK026635 2011 bp mRNA linear PRI 29-SEP-2000

Homo sapiens cDNA: FLJ22982 fis, clone KAT11454, highly similar to AF131235 Homo sapiens N-acetylglyucosamine 6-O-sulfotransferase

MRNA

AK026635

VERSION

AK026635.1 GI:10439531

KEYWORDS

oligo cloning; fis (full insert sequence).

SOURCE

Homo sapiens signal-fing cell carcinoma cell_line:KATO III cDNA to mRNA, clone 11b:KAT clone:KAT11454.

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiindee; Homo.

REFERENCE

1 (sites)

Watanabe,K., Kumagai,A., Itakura,S., Yamazaki,M., Tashiro,H., Ota,T., Suzuki,Y., Ohashi,M., Nishi,T., Shibahara,T., Tanaka,T., Nakamura,Y., Isogai,T. and Sugano,S.

NEO human cDNA sequencing project

unpublished

2 (bases 1 to 2011)

Sugano,S., Suzuki,Y., Ota,T., Ohashi,M., Nishi,T., Isogai,T., Shibahara,T., Tanaka,T. and Nakamura,Y.

Direct Submission

Submitted (29-AUG-2000) Sumio Sugano, Institute of Medical Science, University of Tokyo, Laboratory of Genome Structure Analysis, Human Genome Center, Shirokane-dai, 4-6-1, Minato-ku, Tokyo 108-8639, Japan (E-mail:cdna@i3s.u-tokyo.ac.jp, Tel:81-3-5449-5286, Fax:81-3-5449-5416)

COMMENT

NEO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology; cDNA library construction, 5' - 3' -end one pass sequencing: Department of Virology and Human Genome Center, Institute of Medical Science, University of Tokyo (partly supported by Science and Technology Agency).

FEATURES

source

1. 2011

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="KAT11454"

/cell_line="KATO III"

/cell_type="signal-fing cell carcinoma"

/clone_lib="KAT"

/note="cloning vector pME18SPL3"

1. 2011

/note="highly similar to AF131235 Homo sapiens N-acetylglyucosamine 6-O-sulfotransferase mRNA"

BASE COUNT

482 a 541 c 475 g 513 t

Query Match 91.3%; Score 1856; DB 9; Length 2011;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1906; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 111 GTTACTTTACAGCTTCTCTGGAGGAGAGTGTCTCTCAAGCCCTGTTGCAAGTCTTCC 170

Db 1 GTTACTTTACAGCTTCTCTGGAGGAGAGTGTCTCTCAAGCCCTGTTGCAAGTCTTCC 60

QY 171 ACTTAGACATGCTACTGCTTAATAAATGAAGCTCTCTGTTCTGTTCCCA 230

Db 61 ACTTAGACATGCTACTGCTTAATAAATGAAGCTCTCTGTTCTGTTCCCA 120

QY 231 TGGCCATCTTGGCTCTATCTTCCACATGATACAGCCACACATCAGCTCCCTGTATGA 290

Db 121 TGGCCATCTTGGCTCTATCTTCCACATGATACAGCCACACATCAGCTCCCTGTATGA 180

QY 291 AGGCACAGCCGAGGCGCATGACAGCTGCTGCTCTCTCTGCGGCTGCGCTCTCT 350

Db 181 AGGCACAGCCGAGGCGCATGACAGCTGCTGCTCTCTCTCTGCGGCTGCGCTCTCT 240

QY 351 TTGTGGGAGAGCTTTTGGGAGCACCAGATGTTTCTACTGATGAGCCGCTGAC 410

Db 241 TTGTGGGAGAGCTTTTGGGAGCACCAGATGTTTCTACTGATGAGCCGCTGAC 300

QY 411 ACGTGTGATGACCTTCAAGCAGACACCGCTGATGCTGACATGCTGTGGGATC 470

Db 301 ACGTGTGATGACCTTCAAGCAGACACCGCTGATGCTGACATGCTGTGGGATC 360

QY 471 TGAATAGGGGCGCTTTCTTGTGCGACATGAGAGCTTTGATGCTTCAATGAACTGTG 530

Db 361 TGAATAGGGGCGCTTTCTTGTGCGACATGAGAGCTTTGATGCTTCAATGAACTGTG 420

QY 531 CCCGAGACAGTCCAGCTCTTCAAGTGGAGAACAGCCGCGCTGTGTTCTGACCTG 590

Db 421 CCCGAGACAGTCCAGCTCTTCAAGTGGAGAACAGCCGCGCTGTGTTCTGACCTG 480

QY 591 CCTGTGACATCATCCCAAGATGAATCATCCCGGGGCTCACAGAGCTCTGATGCA 650

Db 481 CCTGTGACATCATCCCAAGATGAATCATCCCGGGGCTCACAGAGCTCTGATGCA 540

QY 651 GTCAACAGCCCTTGTAGAGTGTGAGAGAGCCCTGCGCTCTACAGCCAGCTGTGCTCA 710

Db 541 GTCAACAGCCCTTGTAGAGTGTGAGAGAGCCCTGCGCTCTACAGCCAGCTGTGCTCA 600

QY 711 AGGAGTGGCTTCTTCAACCTGCACTGCTTACCCTGTGTAAGACCCCTCCCTCA 770

Db 601 AGGAGTGGCTTCTTCAACCTGCACTGCTTACCCTGTGTAAGACCCCTCCCTCA 660

QY 771 ACGTGCATCTGTGACAGTGTCCGGAGCCCGCGGCGGTTCGGTCCGAGAACGCA 830

Db 661 ACGTGCATCTGTGACAGTGTCCGGAGCCCGCGGCGGTTCGGTCCGAGAACGCA 720

QY 831 CAAGAGAGATCTCATGATTTGACAGTCCGATTTGATGGGAGCATGAGCAAAACCA 890

Db 721 CAAGAGAGATCTCATGATTTGACAGTCCGATTTGATGGGAGCATGAGCAAAACCA 780

QY 891 AGAAGAGAGACCAACCCCTCTATGTGATGACAGTATGCAAGCCAGCTGAGATCT 950

Db 781 AGAAGAGAGACCAACCCCTCTATGTGATGACAGTATGCAAGCCAGCTGAGATCT 840

QY 951 ACAAGACATCTCAAGTCTTGGCCAAAGGCCCTGACAGAAACCTACCTGTTGCGATAG 1010

Db 841 ACAAGACATCTCAAGTCTTGGCCAAAGGCCCTGACAGAAACCTACCTGTTGCGATAG 900

QY 1011 AGGACCTGGCTGAGCCCTGTTGGCCAGACTTCCCGAATGATATTAATTCGTTGGATTGG 1070

Db 901 AGGACCTGGCTGAGCCCTGTTGGCCAGACTTCCCGAATGATATTAATTCGTTGGATTGG 960

QY 1071 AATCTTGGCCATCTTCAAGCTGGTGCATATCATACCCGAGAGGAGGATGGGATG 1130

Db 961 AATCTTGGCCATCTTCAAGCTGGTGCATATCATACCCGAGAGGAGGATGGGATG 1020

QY 1131 ACCAGCCTTCCACACAAATGCCAGGATGCCCTAATGTCCTCCAGGCTGGCCGTGT 1190
 Db 1021 ACCAGCCTTCCACACAAATGCCAGGATGCCCTAATGTCCTCCAGGCTGGCCGTGT 1080
 QY 1191 CTTTGCCCTATGAAAGGTTCTGCATCTTCAAAAGCCTGGGCAATGCCATGAAATTTGC 1250
 Db 1081 CTTTGCCCTATGAAAGGTTCTGCATCTTCAAAAGCCTGGGCAATGCCATGAAATTTGC 1140
 QY 1251 TGGGCTACCGCCACGCTGAGATCTGAACAAGAGAGAAAGCCTGGTGCATCTTCTGT 1310
 Db 1141 TGGGCTACCGCCACGCTGAGATCTGAACAAGAGAGAAAGCCTGGTGCATCTTCTGT 1200
 QY 1311 CTACTGAGACTGCTCCCTGAGCAAACTCCTTAAGAGGTTGAGAGGCTTTGCTGCCACT 1370
 Db 1201 CTACTGAGACTGCTCCCTGAGCAAACTCCTTAAGAGGTTGAGAGGCTTTGCTGCCACT 1260
 QY 1371 GGTGTCAGCCTCAGTCACTTTCTCTGTAATGCTTGTGAGCCTTGGCTACATCTGAGGCT 1430
 Db 1261 GGTGTCAGCCTCAGTCACTTTCTCTGTAATGCTTGTGAGCCTTGGCTACATCTGAGGCT 1320
 QY 1431 TAATTAATGTCGTGCTGATTCACACTGAGTGTGATTTGTGCAACAGCTGCTCAAGCAG 1490
 Db 1321 TAATTAATGTCGTGCTGATTCACACTGAGTGTGATTTGTGCAACAGCTGCTCAAGCAG 1380
 QY 1491 AAGGACTTTGTGTCATGCTTGTGCTAGAAACAGACAGCTGGGACCTTATGTGAGCAG 1550
 Db 1381 AAGGACTTTGTGTCATGCTTGTGCTAGAAACAGACAGCTGGGACCTTATGTGAGCAG 1440
 QY 1551 CACATCCACAGTGAAGAAACAGGATATGCTTCTTCTTCTGATCTTCTGCTGCTGCTG 1610
 Db 1441 CACATCCACAGTGAAGAAACAGGATATGCTTCTTCTTCTGATCTTCTGCTGCTGCTG 1500
 QY 1611 CAGACTTCAGAGACTTTGTGGCCTGGAGGCTTATTAAGCAGACAGTACAGTGAAT 1670
 Db 1501 CAGACTTCAGAGACTTTGTGGCCTGGAGGCTTATTAAGCAGACAGTACAGTGAAT 1560
 QY 1671 TGATCCATTAACCTCCCTGTCACATCTTGCCCAATGGGAGATGATTTTACCAAGA 1730
 Db 1561 TGATCCATTAACCTCCCTGTCACATCTTGCCCAATGGGAGATGATTTTACCAAGA 1620
 QY 1731 GCTACCCAGCATTTTCCACAGATGCAAAATTCGAGCCCTTGAGAGTCCAGTGAATC 1790
 Db 1621 GCTACCCAGCATTTTCCACAGATGCAAAATTCGAGCCCTTGAGAGTCCAGTGAATC 1680
 QY 1791 AAGGAGAGATGGGAGCAAGGTGGATGCTTACTTAAGGCTGAGCATCAGCATATC 1850
 Db 1681 AAGGAGAGATGGGAGCAAGGTGGATGCTTACTTAAGGCTGAGCATCAGCATATC 1740
 QY 1851 GGTATATCAGAAATATGAAACAAATCTCTGCACAAAGAGCAGCTTAAAGTACAGG 1910
 Db 1741 GGTATATCAGAAATATGAAACAAATCTCTGCACAAAGAGCAGCTTAAAGTACAGG 1800
 QY 1911 GTGCTGGGGCGCATTTGAATATCATCTCCCTGCTGATTTTCCCATCAGATAGAACT 1970
 Db 1801 GTGCTGGGGCGCATTTGAATATCATCTCCCTGCTGATTTTCCCATCAGATAGAACT 1860
 QY 1971 TTGACCTGTGAAGCTGCCATCTGTTAATACTAAATTTCCCAATTAAG 2017
 Db 1861 TTGACCTGTGAAGCTGCCATCTGTTAATACTAAATTTCCCAATTAAG 1907
 RESULT 4
 AC010547 183228 bp DNA linear PRI 29-SEP-2001
 LOCUS Homo sapiens chromosome 16 clone RP11-510M2, complete sequence.
 DEFINITION AC010547
 ACCESSION AC010547
 VERSION AC010547.9 GI:15808510
 KEYWORDS HTG.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 REFERENCE 1 (bases 1 to 183228)

AUTHORS DOE Joint Genome Institute.
 TITLE Sequencing of Human Chromosome 16
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 183228)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 3 (bases 1 to 183228)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (29-SEP-2001) Production Sequencing Facility, DOE Joint
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 On Sep 29, 2001 this sequence version replaced gi:14589436.
 COMMENT This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 Genbank flat file format but are available as part
 of this entry's ASN.1 file.

 Sequence Quality Assessment:
 This entry has been annotated with sequence quality
 estimates computed by the Phrap assembly program.
 All manually edited bases have been reduced to quality zero.
 Quality levels above 40 are expected to have less than
 1 error in 10,000 bp.
 Base-by-base quality values are not generally visible from the
 Genbank flat file format but are available as part
 of this entry's ASN.1 file.

 FEATURES
 source 1.183228
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /clone="RP11-510M2"
 BASE COUNT 51901 a 42098 c 40026 g 49203 t
 ORIGIN
 Query Match 88.9%; Score 1806; DB 9; Length 183228;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1856; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 161 AAGTCTTCACCTCAGCAAAATGCTACTGCTTAAATTAAGTCTGCTGTTTCTG 220
 Db 142584 AAGTCTTCACCTCAGCAAAATGCTACTGCTTAAATTAAGTCTGCTGTTTCTG 142643
 QY 221 GTTTCCAGATGGCCATCTTGCTCTATTTCTTCCATGATGACGCCAACAATCAGCTCG 280
 Db 142644 GTTTCAGATGGCCATCTTGCTCTATTTCTTCCATGATGACGCCAACAATCAGCTCG 142703
 QY 281 CTGCTATGAAGGACAGCCCGGCGGAGTGCAGTGTCTGTCTTCTGCGCTCT 340
 Db 142704 CTGCTATGAAGGACAGCCCGGCGGAGTGCAGTGTCTGTCTTCTGCGCTCT 142763
 QY 341 GGCCTCTCTTTTGTGGGAGCTTTTGGGACAGCACCAGATTTTCTACCTGATGAG 400
 Db 142764 GGCCTCTCTTTTGTGGGAGCTTTTGGGACAGCACCAGATTTTCTACCTGATGAG 142823
 QY 401 CCGGCTGCGAGCTGTGATGACCTTCAAGCAGCAGCCGCTGAGTGTGACATGCT 460
 Db 142824 CCGGCTGCGAGCTGTGATGACCTTCAAGCAGCAGCCGCTGAGTGTGACATGCT 142883
 QY 461 GTGCGGATCTGATAGGCGCGCTCTTGTGGAGATGAGGCTTTGATGCTATAGT 520
 Db 142884 GTGCGGATCTGATAGGCGCGCTCTTGTGGAGATGAGGCTTTGATGCTATAGT 142943
 QY 521 GAACCTGTGCTCCCGGAGACATGCTTCACTGCTTCACTGGGAGAACCGCGGCTGTGT 580

OY	1661	TCACGTGCATTTGANCCTAAACCTCCCGTCCACATCTTGCCCAATGGGAAGATCCT	1720
OY	1661		1720
OY	1661		1720
OY	1661		1720
Db	144084	TCAATGGATTGATGCCATAAACCTCCCGTCCACATCTTGCCCAATGGGAAGATCCT	144143
OY	1721	TCACCAGAAGACTCACACGCAATTTTCCACAAGATGCAAATTTCTGAGCCCTTGGAAGTTCC	1780
Db	144144	TCACCAGAAGACTCACACGCAATTTTCCACAAGATGCAAATTTCTGAGCCCTTGGAAGTTCC	144203
OY	1781	CAGTGGATTCAAGNAGAAGATGGGAACAAGTTGGATGGATCTTATGAGCTTGACCAT	1840
Db	144204	CAGTGGATTCAAGNAGAAGATGGGAACAAGTTGGATGGATCTTATGAGCTTGACCAT	144263
OY	1841	CACACGCTATCGGTAAATCATCAAAAATATGAAACAAAATCTCTGCACAAAAAAGCAAGCTCTTA	1900
Db	144264	CACACGCTATCGGTAAATCATCAAAAATATGAAACAAAATCTCTGCACAAAAAAGCAAGCTCTTA	144323
OY	1901	AGTTACACAGGGGCGCTGGGCGTCGATTTAAATATCATCTCCCGCTCGCATTTTCCATCAC	1960
Db	144324	AGTTACACAGGGGCGCTGGGCGTCGATTTAAATATCATCTCCCGCTCGCATTTTCCATCAC	144383
OY	1961	ATAGAAGACTTTGACCTGTGAAGCTGCGATCTGTTAAATACATTAANTTCCCANAATPAG	2017
Db	144384	ATAGAAGACTTTGACCTGTGAAGCTGCGATCTGTTAAATACATTAANTTCCCANAATPAG	144440
RESULT 5			
AF280088		1992 bp mRNA linear PRI 20-FEB-2001	
LOCUS		Homo sapiens I-selectin ligand sulfotransferase GST-3 mRNA,	
DEFINITION		complete cds.	
ACCESSION		AF280088	
VERSION		AF280088.1 GI:12060807	
KEYWORDS			
SOURCE			
ORGANISM		Homo sapiens.	
REFERENCE		Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
AUTHORS		1 (bases 1 to 1992)	
		Hemmerich,S., Lee,J.K., Bhakta,S., Bistrup,A., Ruddle,N.R. and Rosen,S.D.	
TITLE		Chromosomal localization and genomic organization for the galactose/N-acetylgalactosamine/N-acetylglucosamine 6-O-sulfotransferase gene family	
JOURNAL		Glycobiology 11 (1), 75-87 (2001)	
MEDLINE		21096027	
PUBMED		11181564	
REFERENCE		2 (bases 1 to 1992)	
AUTHORS		Hemmerich,S., Bhakta,S., Lee,J.-K., Bistrup,A., Ruddle,N.R. and Rosen,S.D.	
TITLE		Direct Submision Submitted (20-JUN-2000) Respiratory Diseases, Roche Bioscience, 3401 Hillview Ave., Palo Alto, CA 94304, USA	
JOURNAL		Location/Qualifiers	
FEATURES			
source			
		1..1992	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="16"	
		/map="16q23.1-q23.2"	
		/cell_type="high endothelial"	
		/tissue_type="tonsil"	
		1..1992	
		/gene="GST3"	
		122..1282	
		/note="similar to Homo sapiens sequence encoded by Genbank Accession Number AF131235"	
		/codon_start=1	
		/product="I-selectin ligand sulfotransferase GST-3"	
		/protein_id="AAG48246.1"	
		/db_xref="gi:12060808"	
		/translation="MLDPRKKMLLFLVSONAILALFFPHYSHTISLSMKRQPERMT VLVISWMSGSFVGQLFGQHPDVYLEMPAHVMVFKEOSTAMVLHMAVRDLIRAVP	
gene			
CDs			

LCDSVPDAMWEPGRROSSIFOMENSRALCSAPACDIIPDEITIPRAHCLLCSOP
PEVEKAKRSTSHVYLKREFFENFLOSLYPLKIDSLNHYLVHPDPAVRSEERK
GDMIDSRIVMG0HQKLEKEDQYVQVQSOQLKYEKTIQSLPKJLORLYLVRY
EDLAAPVATSRMYEKEFEFLPHLOTWVNIIRGKMGMHAHTNADALNYSQAW
RWSLPYEVYSLOKACGDANMLIGRYHRSBOQRNLLDLSTWYVEQJH"
BASE COUNT 482 a 540 c 489 g 481 t
ORIGIN

Query Match 86.4%; Score 1756; DB 9; Length 1992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1856; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 160 CAGGCTTCCACTTCAGACATGCTGCTTAAAAAATGANGCTCCGCTGCTTCT 219
DB 100 CAGGCTTCCACTTCAGACATGCTGCTTAAAAAATGANGCTCCGCTGCTTCT 159
QY 220 GGTTCGCCAGATGGCCATCTGGCTCTATCTTCACATGTACAGCCACAACTCAGCTC 279
DB 160 GGTTCGCCAGATGGCCATCTGGCTCTATCTTCACATGTACAGCCACAACTCAGCTC 219
QY 280 CCGTCTATGAGGAGCAGACCCCGAGCATGACGTGCTGCTGCTCTCTGCGCTC 339
DB 220 CCGTCTATGAGGAGCAGACCCCGAGCATGACGTGCTGCTGCTCTCTGCGCTC 279
QY 340 TGGCTCTCTTTTGTGGGCGACCTTTTGGCGAGCACCAGATGTTTCTACCTGATGA 399
DB 280 TGGCTCTCTTTTGTGGGCGACCTTTTGGCGAGCACCAGATGTTTCTACCTGATGA 339
QY 400 GCGCCCTGGCAGCTGTGATACCTTCAGAGAGCAGCCGCTGATGTGACATGGC 459
DB 340 GCGCCCTGGCAGCTGTGATACCTTCAGAGAGCAGCCGCTGATGTGACATGGC 399
QY 460 TGTGGGAGATGATACGGGCGCTCTTGTGCACATGACGCTTTGATGCTTACAT 519
DB 400 TGTGGGAGATGATACGGGCGCTCTTGTGCACATGACGCTTTGATGCTTACAT 459
QY 520 GGAACCTGTGCTCCCGAGACAGTCCAGCTCTTTCAGTGGGAGAAAGCCGGGCTGTG 579
DB 460 GGAACCTGTGCTCCCGAGAGCAGTCCAGCTCTTTCAGTGGGAGAAAGCCGGGCTGTG 519
QY 580 TTCTGCACCTGCTGTGACATCATCCCAAGATGAATTCACCCGGGCTCACTGAG 639
DB 520 TTCTGCACCTGCTGTGACATCATCCCAAGATGAATTCACCCGGGCTCACTGAG 579
QY 640 GCTCTGTGACATCAAGCCCTTTAGAGTGTGAGAGAGGCTGCGCTCTTACAGCA 699
DB 580 GCTCTGTGACATCAAGCCCTTTAGAGTGTGAGAGAGGCTGCGCTCTTACAGCA 639
QY 700 CGTGTGTCAGAGAGGAGGCGCTTCTGACCTGCACTGCTCTACCCGCTGCTGAAGA 759
DB 640 CGTGTGTCAGAGAGGAGGCGCTTCTGACCTGCACTGCTCTACCCGCTGCTGAAGA 699
QY 760 CCGCTCCCTCAACCTGATATGTCGACCTGATCCGGAGCCCGCGGCGTGTCCGCTC 819
DB 700 CCGCTCCCTCAACCTGATATGTCGACCTGATCCGGAGCCCGCGGCGTGTCCGCTC 759
QY 820 CCGAAGAGCAGAAAGGAGATCTATATGACAGTGCATTTGATGGGGCAGCAGTA 879
DB 760 CCGAAGAGCAGAAAGGAGATCTATATGACAGTGCATTTGATGGGGCAGCAGTA 819
QY 880 GCAGAACTCAAGAGAGAGCAGCAACCTACTATGTGATGAGAGTCATCTGCCAAGCA 939
DB 820 GCAGAACTCAAGAGAGAGCAGCAACCTACTATGTGATGAGAGTCATCTGCCAAGCA 879
QY 940 GCTGAGATCTACAGACATCCAGTCTTGTGCCAAGGCGCTGACAGAGCCTTACTGCT 999
DB 880 GCTGAGATCTACAGACATCCAGTCTTGTGCCAAGGCGCTGACAGAGCCTTACTGCT 939
QY 1000 TGTGGCATGAGAGCTGTGCTGAGCCCTGTGAGCCAGATTCGCGAATGTATGAT 1059
DB 940 TGTGGCATGAGAGCTGTGCTGAGCCCTGTGAGCCAGATTCGCGAATGTATGAT 999
QY 1060 CGTGGAGATGGAATTTCTTCCCATCTTTCAGACCTGGGTGATTAACATCACCCGAGCA 1119

DB 1000 CGTGGAGATGGAATTTCTTCCCATCTTTCAGACCTGGGTGATTAACATCACCCGAGCA 1059
QY 1120 GGGCATGGGTGACACAGCTTTTCCACAGAAATGCCAGGATGCCCTTAATGCTCCAGGC 1179
DB 1060 GGGCATGGGTGACACAGCTTTTCCACAGAAATGCCAGGATGCCCTTAATGCTCCAGGC 1119
QY 1180 TTGGCGCTGTCCTTTGCCCTATGAAAAGTTTTCGAGTTCAGAAAAGCCTTGCGCATGC 1239
DB 1120 TTGGCGCTGTCCTTTGCCCTATGAAAAGTTTTCGAGTTCAGAAAAGCCTTGCGCATGC 1179
QY 1240 CATGATTTTGGTGGCTACCGCCCAAGTCAGATGACAGAAACAGAAACCTGTGCT 1299
DB 1180 CATGATTTTGGTGGCTACCGCCCAAGTCAGATGACAGAAACAGAAACCTGTGCT 1239
QY 1300 GGATCTTCTGCTACCTGAGTGTCCCTGAGCAATCCACTAAGAGGTTGAGAGGCT 1359
DB 1240 GGATCTTCTGCTACCTGAGTGTCCCTGAGCAATCCACTAAGAGGTTGAGAGGCT 1299
QY 1360 TGTGCCACCTGTGTGACGCTCACTCTTCTGATGCTTCTAGCCTTGCTTACA 1419
DB 1300 TGTGCCACCTGTGTGACGCTCACTCTTCTGATGCTTCTAGCCTTGCTTACA 1359
QY 1420 TCTGAGACCTTTAATACATGTCGTGGGTATACACTGAGTGTGAGTGTGCTACAG 1479
DB 1360 TCTGAGACCTTTAATACATGTCGTGGGTATACACTGAGTGTGAGTGTGCTACAG 1419
QY 1480 TGTCTAACAGAGAGCTTTTGTGCTCATGCTTGTGTACAAAAACAGACTGGGAACT 1539
DB 1420 TGTCTAACAGAGAGCTTTTGTGCTCATGCTTGTGTACAAAAACAGACTGGGAACT 1479
QY 1540 TATGTGACAGACATCCACAGTGAAGAGGATTTGCTCTCTCTTCTTGTACT 1599
DB 1480 TATGTGACAGACATCCACAGTGAAGAGGATTTGCTCTCTCTTCTTGTACT 1539
QY 1600 TCTGTGTGGGAGACTTCAGAGACTTTGTGGCCCTGAGAGGCTTTAAGACAGCAGACT 1659
DB 1540 TCTGTGTGGGAGACTTCAGAGACTTTGTGGCCCTGAGAGGCTTTAAGACAGCAGACT 1599
QY 1660 ATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGCCCAATGGGGAATGATCT 1719
DB 1600 ATCAGTGAATGATTCATTAACCTCCCTGTCACATCTTGCCCAATGGGGAATGATCT 1659
QY 1720 TTCAACAAAGAGCTCACAGCAATTTCCACAGAGATCCAATTTGAGCCCTTGAATTC 1779
DB 1660 TTCAACAAAGAGCTCACAGCAATTTCCACAGAGATCCAATTTGAGCCCTTGAATTC 1719
QY 1780 CCAATGATTTAAGAGAGAAAGTGGAAACAAGTTGATGCTTACTATGAGCTTGACCA 1839
DB 1720 CCAATGATTTAAGAGAGAAAGTGGAAACAAGTTGATGCTTACTATGAGCTTGACCA 1779
QY 1840 TCACAGCTATGCTATACAGAAATATGAAACAAATCTGACACAAAAGCAAGCTCT 1899
DB 1780 TCACAGCTATGCTATACAGAAATATGAAACAAATCTGACACAAAAGCAAGCTCT 1839
QY 1900 AAGTTCAAGGGGTGCTGGGCTGCATTTGAATATCACTTCCCTGCTGATTTCCATCA 1959
DB 1840 AAGTTCAAGGGGTGCTGGGCTGCATTTGAATATCACTTCCCTGCTGATTTCCATCA 1899
QY 1960 CATAGAGACTTTGACCTGTAAGCTGCCATCTGTAATCTAAATTTCCCAATAG 2017
DB 1900 CATAGAGACTTTGACCTGTAAGCTGCCATCTGTAATCTAAATTTCCCAATAG 1957

RESULT 6
AF149783
LOCUS AF149783 1333 bp mRNA linear PRI 02-JUL-2001
DEFINITION Homo sapiens L-selectin ligand sulfotransferase mRNA, complete cds.
ACCESSION AF149783
VERSION AF149783.1 GI:13897503
KEYWORDS
SOURCE
ORGANISM Homo sapiens

Db	552	CTCCTGTGCAATCAACAGCCCTTTGAGGTGGTGGAGAAGGCGCTCCCGCTCTTACAGCCAC	611
QY	701	GTGGTGTGCAGAGAGGAGCGCTTCTTCAACCTGGAGTCCCTCTCAACCGCTCTGAAGAG	760
Db	612	GTGGTGTGCAGAGAGGAGCGCTTCTTCAACCTGGAGTCCCTCTCAACCGCTCTGAAGAG	671
QY	761	CCCTCCCTCAACCTGCATATGTGACACCTGGTCCGGAGCCCGCGGGCGTTCCGTTCC	820
Db	672	CCCTCCCTCAACCTGCATATGTGACACCTGGTCCGGAGCCCGCGGGCGTTCCGTTCC	731
QY	821	CGAAGACGCAAAAGGGAGATCTCATGATTACAGTGTGCATGTGGATGGGACACATGAG	880
Db	732	CGAAGACGCAAAAGGGAGATCTCATGATTACAGTGTGCATGTGGATGGGACACATGAG	791
QY	881	CAGAAACTCAAGAGAGAGAGCAACCCCTACTATGTGATGACGGTCACTGTGCCAAAGCCAG	940
Db	792	CAAAAACTCAGAGAGAGAGCAACCCCTACTATGTGATGACGGTCACTGTGCCAAAGCCAG	851
QY	941	CTGGAGATCTCAAGAGACCATCAAGTCTTGGCCCAAGGCCCTTGACAGAAAGCTTACTGCTT	1000
Db	852	CTGGAGATCTCAAGAGACCATCAAGTCTTGGCCCAAGGCCCTTGACAGAAAGCTTACTGCTT	911
QY	1001	GTGGGCTATGAGAGACCTGGCTCGAGCCCTGTGGGCCAGACTTCCTCCGAATGATGATTC	1060
Db	912	GTGGGCTATGAGAGACCTGGCTCGAGCCCTGTGGGCCAGACTTCCTCCGAATGATGATTC	971
QY	1061	GTGGGATGGAATCTTGGCCCATCTTTCAGACCTGGGTGATTAACATCACCCGAGGCAAG	1120
Db	972	GTGGGATGGAATCTTGGCCCATCTTTCAGACCTGGGTGATTAACATCACCCGAGGCAAG	1031
QY	1121	GGCATGGGTGACACACGCTTTCACACAAATGCGAGGATGGCCCTTAATGTCTCCAGGCT	1180
Db	1032	GGCATGGGTGACACACGCTTTCACACAAATGCGAGGATGGCCCTTAATGTCTCCAGGCT	1091
QY	1181	TGGCGCTGGTCTTGGCCCTATGAAAAGGTCTTTCGACTTCAGAAAAGCCCTGTGGGATGCC	1240
Db	1092	TGGCGCTGGTCTTGGCCCTATGAAAAGGTCTTTCGACTTCAGAAAAGCCCTGTGGGATGCC	1151
QY	1241	ATGAATTTGCTGGGCTACCGCCACGCTCAGATCTGATAACAAGAACAGAAACCTGTTGCTG	1300
Db	1152	ATGAATTTGCTGGGCTACCGCCACGCTCAGATCTGATAACAAGAACAGAAACCTGTTGCTG	1211
QY	1301	GATCTCTGTCTACTACCTGAGTGTCCCTGAGCAAAATCCACTAAGAGGATTGAGAAAGCTTT	1360
Db	1212	GATCTCTGTCTACTACCTGAGTGTCCCTGAGCAAAATCCACTAAGAGGATTGAGAAAGCTTT	1271
QY	1361	GCTGCCACCTGGTGTGACGCTCAGCTCAGTCTCTGTAATGCTTCTGAGCCTTGCCTACAT	1420
Db	1272	GCTGCCACCTGGTGTGACGCTCAGCTCAGTCTCTGTAATGCTTCTGAGCCTTGCCTACAT	1331
QY	1421	CT 1422	
Db	1332	CT 1333	
RESULT 7			
LOCUS	AX381256	517 bp	DNA
DEFINITION	Sequence 194 from Patent WO0212280.	Linear	PAT 18-MAR-2002
ACCESSION	AX381256		
VERSION	AX381256.1	GI:19576075	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Pyle,R.A., Xu,J. and Secrist,H.		
1	Compositions and methods for the therapy and diagnosis of colon		
JOURNAL	cancer		
FEATURES	Patient: WO 0212280-A 194 14-FEB-2002;		
	CORIXA CORPORATION (US)		
	Location/Qualifiers		

```
source 1..517
        /db_xref="taxon:9606"
BASE COUNT 121 a 130 c 124 g 142 t
ORIGIN
Query Match 24.9%; Score 505; DB 6; Length 517;
Best Local Similarity 100.0%; Pred. No. 5,6e-290;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 GGCCTTGGCCACCGCTGGGTAGCCTCAGTCAGTCACTTCTCTAAATGCTTCTAGCCTTGC 1414
    |||||
DB 13 GGCCTTGGCCACCGCTGGGTAGCCTCAGTCAGTCACTTCTCTAAATGCTTCTAGCCTTGC 72
    |||||
QY 1415 CTACATCTCTGAGCCTTAATCATATGTCGTGGGTATCAGCTAGTGTGAGTGTGCTC 1474
    |||||
DB 73 CTACATCTCTGAGCCTTAATCATATGTCGTGGGTATCAGCTAGTGTGAGTGTGCTC 132
    |||||
QY 1475 AACGCTGCTCAAGCAGAGAGGACTTTGTGTCATGCTGTGTCCTAGAAAAACAGACTGGGG 1534
    |||||
DB 133 AACGCTGCTCAAGCAGAGAGGACTTTGTGTCATGCTGTGTCCTAGAAAAACAGACTGGGG 192
    |||||
QY 1535 AACCTATGTCAGCAGACATCCACAGGATGAAACAGGATGCTTCTCTTCTTCTTCT 1594
    |||||
DB 193 AACCTATGTCAGCAGACATCCACAGGATGAAACAGGATGCTTCTCTTCTTCTTCTT 252
    |||||
QY 1595 GATCTTCTCTGCTGGGACACTTTCAGAGACTTGTGGCCTGAGGCTTATTAGCAGCAGC 1654
    |||||
DB 253 GATCTTCTCTGCTGGGACACTTTCAGAGACTTGTGGCCTGAGGCTTATTAGCAGCAGC 312
    |||||
QY 1655 ACAGTATGTCAGGAAATGATTCATTAACCTCCCTGTCACATCTGCCCCAATGGGGAGT 1714
    |||||
DB 313 ACAGTATGTCAGGAAATGATTCATTAACCTCCCTGTCACATCTGCCCCAATGGGGAGT 372
    |||||
QY 1715 GATCTTTCACCAAGAGCTCACACCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGG 1774
    |||||
DB 373 GATCTTTCACCAAGAGCTCACACCATTTTCCACAGAGATGCAATTTCTGAGCCCTTGG 432
    |||||
QY 1775 AGTTCCTCAGTATTCAGAGAGAGAGTGGGACAAAGGTGATGCTTACTTATAGCTT 1834
    |||||
DB 433 AGTTCCTCAGTATTCAGAGAGAGAGTGGGACAAAGGTGATGCTTACTTATAGCTT 492
    |||||
QY 1835 GACCATTCACAGCTATCGTATTCAG 1859
    |||||
DB 493 GACCATTCACAGCTATCGTATTCAG 517
    |||||

RESULT 8
AX381326 389 bp DNA linear PAT 18-MAR-2002
LOCUS AX381326
DEFINITION Sequence 264 from Patent WO0212280.
ACCESSION AX381326
VERSION AX381326.1 GI:19576145
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Pyle, R.A., Xu, J. and Secrist, H.
TITLE Compositions and methods for the therapy and diagnosis of colon
        Cancer
JOURNAL Patent: WO 0212280-A 264 14-FEB-2002;
CORIXA CORPORATION (US)
FEATURES
        source 1..389
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
BASE COUNT 121 a 83 c 76 g 109 t
ORIGIN
Query Match 14.1%; Score 286; DB 6; Length 389;
Best Local Similarity 100.0%; Pred. No. 9.2e-159;

Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1732 CTCACACGATTTTCCACAGAGATGCAAAATTCGTAGCCCTTGGAGTCCAGTGATTC 1791
    |||||
DB 13 CTCACACGATTTTCCACAGAGATGCAAAATTCGTAGCCCTTGGAGTCCAGTGATTC 72
    |||||
QY 1792 AGGAAGAGAGGAGGAGAGAGTGGATGCTTACTATGAGCTTGACATCAGCTATTCG 1851
    |||||
DB 73 AGGAAGAGAGGAGGAGAGAGTGGATGCTTACTATGAGCTTGACATCAGCTATTCG 132
    |||||
QY 1852 GTATCAGAAATATGAAACAAATCTCTGCACAAAAGAGAGAGCTTTAGTTCACAGG 1911
    |||||
DB 133 GTATCAGAAATATGAAACAAATCTCTGCACAAAAGAGAGAGCTTTAGTTCACAGG 192
    |||||
QY 1912 TGCCCTGGGCTGCATTTGAAATTCATCTCCCTCTGCATTTTCCATTCATCAGAGACT 1971
    |||||
DB 193 TGCCCTGGGCTGCATTTGAAATTCATCTCCCTCTGCATTTTCCATTCATCAGAGACT 252
    |||||
QY 1972 TGACCTGTGAAGCTGCCATCTGTATTAATCTAAATTCACAAATTAAG 2017
    |||||
DB 253 TGACCTGTGAAGCTGCCATCTGTATTAATCTAAATTCACAAATTAAG 298
    |||||

RESULT 9
AF176839
LOCUS AF176839
DEFINITION Homo sapiens intestine N-acetylglucosamine 6-O-sulfotransferase
        (I-GlcNAc-6-ST) gene, complete cds.
ACCESSION AF176839
VERSION AF176839.1 GI:5917707
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1
AUTHORS Lee, J.K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
TITLE Cloning and characterization of a mammalian
        N-acetylglucosamine-6-sulfotransferase that is highly restricted to
        intestinal tissue
JOURNAL Biochem. Res. Commun. 263 (2), 543-549 (1999)
MEDLINE 99423499
PUBMED 10491328
REFERENCE 2 (bases 1 to 1462)
AUTHORS Lee, J.-K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
TITLE Direct Submission
        Submitted (10-AUG-1999) Respiratory Diseases, Roche Bioscience,
        3401 Hillview Avenue, Palo Alto, CA 94304, USA
FEATURES
        source 1..1462
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /chromosome="16"
                /map="16q23.1"
                /tissue_type="intestine"
                /note="possibly involved in biosynthesis of mucosal
                ligands for L-selectin"
                /codon_start=1
                /product="N-acetylglucosamine 6-O-sulfotransferase"
                /protein_id="A056001.1"
                /db_xref="GI:5917708"
                /translation="MWLPRESKTVTVTLAQTCLLPIISRPSPAGGDRVHV
                LVLSWMSGSSFLGOLFSDPVEFLMEPAWHVWTLTSGSAAATLMAVRDLRSIFL
                CMDVDPAWPOSRLNSAFPNWATSRLCSPPKCSAFPRTISKQVCTILCTROPFS
                LAREACHRSYHVLEKRFENLDVLPPLSDPALNRIYHLVDPAVAVSRRAAGPI
                LARDNGIVIGTNGKWEADPHRLIREVCRSHVRIAEATLKPPPLRGYRIVREED
```

BASE COUNT 239 a 510 c 446 g 267 t
 ORIGIN
 Query Match 2.9%; Score 59; DB 9; Length 1462;
 Best Local Similarity 100.0%; Pred. No. 1e-22;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 736
 |||||||
 Db 719 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 777

RESULT 10
 AX327330 1647 bp DNA linear PAT 07-JAN-2002
 LOCUS Sequence 15 from Patent WO0179468.
 ACCESSION AX327330
 VERSION AX327330.1 GI:18097876
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 Policky, J.L., Hafalia, A., Burford, N., Ring, H.Z., Lal, P.,
 Tribouley, C.M., Yao, M.G., Yue, H., Tang, Y.T., Patterson, C., Das, D.,
 Sanjanwala, M.S., Gandhi, A.R., Reddy, R., Khan, F.A., Baughn, M.R.,
 Ramkumar, J., Griffin, J.A. and Au-Young, J.
 TITLE Drug metabolizing enzymes
 JOURNAL Patent: WO 0179468-A 15 25-OCT-2001;
 Incyte Genomics, Inc. (US)

FEATURES
 source
 1..1647
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /note="Incyte ID No: 7472777CB1"
 BASE COUNT 287 a 560 c 501 g 299 t
 ORIGIN

Query Match 2.9%; Score 59; DB 6; Length 1647;
 Best Local Similarity 100.0%; Pred. No. 1e-22;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 736
 |||||||
 Db 488 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 546

RESULT 11
 AF176838 2170 bp mRNA linear PRI 22-SEP-1999
 LOCUS
 DEFINITION Homo sapiens intestine N-acetylglucosamine 6-O-sulfotransferase
 (I-G1CNAc-6-ST) mRNA, complete cds.
 ACCESSION AF176838
 VERSION AF176838.1 GI:5917705
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2170)
 Lee, J.K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
 TITLE Cloning and characterization of a mammalian
 N-acetylglucosamine-6-sulfotransferase that is highly restricted to
 intestinal tissue
 JOURNAL Biochem. Biophys. Res. Commun. 263 (2), 543-549 (1999)
 MEDLINE 99423499
 PUBMED 10491328
 REFERENCE 2 (bases 1 to 2170)
 Lee, J.-K., Bhakta, S., Rosen, S.D. and Hemmerich, S.
 TITLE Direct Submission

JOURNAL Submitted (10-AUG-1999) Respiratory Diseases, Roche Bioscience,
 3401 Hillview Avenue, Palo Alto, CA 94304, USA
 FEATURES
 source
 1..2170
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16q23.1"
 /tissue_type="Intestine"
 1..2170
 /gene="I-G1CNAc-6-ST"
 344..1516
 /gene="I-G1CNAc-6-ST"
 /note="possibly involved in biosynthesis of mucosal
 ligands for L-selectin"
 /codon_start=1
 /product="N-acetylglucosamine 6-O-sulfotransferase"
 /protein_id="AAB56000.1"
 /db_xref="GI:5917706"

translation="MWLPRFSKTVTVLLAQTCLLFTIISRPSPSSPAGEEDRVH
 LVLSWRGSSFLGOLFQSDHVDVFLMEPMHWTLTSGSAATLHMAVADLRSLFL
 CDMDYFDYMPQSNLSAFTFMATSRALCSRPACSAFPRGTISKDYCKRLCTRPFS
 LARACRSYSHVLEKEVFNLQVLPULSDPALNLKIVLVRDRAVLRSRAGFI
 LARNGIVLGTNGKWEADPHLRIVRCRSHVRIABAATLKPPPLRGYRIWFED
 LARPLAEIRALVAFGLTLPQLEAMVHNTHSGSGIKPIEAFHSSRNARVSCAM
 RHALPFTKILRVQECAGALQLIGRPVYSADQORDLTLVLRCPDHFSSWASPD"

BASE COUNT 449 a 682 c 612 g 427 t
 ORIGIN
 Query Match 2.9%; Score 59; DB 9; Length 2170;
 Best Local Similarity 100.0%; Pred. No. 1e-22;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 736
 |||||||
 Db 834 AGGCGTCCCGCTCTACAGCCAGCGTGTGCTCAAGAGGTGGCTCTTCAACCTGCAG 892

RESULT 12
 AF219990 2544 bp mRNA linear PRI 26-OCT-2000
 LOCUS
 DEFINITION Homo sapiens corneal N-acetylglucosamine-6-O-sulfotransferase
 (CHST6) mRNA, complete cds.
 ACCESSION AF219990
 VERSION AF219990.1 GI:11023145
 KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 2544)
 Akama, T.O., Nishida, K., Nakayama, J., Watanabe, H., Fujiwara, T.,
 Nakamura, T., Dote, A., Kawasaki, S., Inoue, Y., Maeda, N., Yamamoto, S.,
 Ozaki, K., Kinoshita, S., Shimomura, Y., Tanigami, A. and Fukuda, M.N.
 TITLE Macular corneal dystrophy type I and type II are caused by distinct
 mutations in a new sulphotransferase gene
 JOURNAL Nat. Genet. 26 (2), 237-241 (2000)
 MEDLINE 20472350
 PUBMED 11017086
 REFERENCE 2 (bases 1 to 2544)
 Akama, T.O. and Fukuda, M.N.
 TITLE Direct Submission
 JOURNAL Submitted (29-DEC-1999) Glycobiology Program, The Burnham
 Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
 source
 1..2544
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="16"
 /map="16q22"
 1..2544
 /gene="CHST6"
 /note="responsible for macular corneal dystrophy"

CDS

693..1880
/gene="CHST6"
/function="Transfers sulfate to 6-OH of GlcNAc in keratan sulfate"
/note="C-GlcNAc6S7"
/codon_start=1
/product="corneal N-acetylglucosamine-6-O-sulfotransferase"
/protein_id="AA026325.1"
/db_xref="GI:11023146"
/translation="MGLPVSSTAVTALLAQTELLFLVSRPSSPAGGEARVETLVLSWRSSSSGQDLFNOHPDYFLMEPAMHWTTLSQSNATLHMAVRDVRSEFLCDMDVDALPWRNLSDLEQMAVRALCPSPACSPAGATISSEAVCKPACSEFLAREACRSYSHVLEKEVRPNLOVYPLISDPALNLRVHLVDRDRAVLRSEQKALARDNGVIGTGTVEADPCGLRVRYRCRSRVRAENATLKPPFLGRVRLVEEDLAREPLAEIRALVARTGSLTLPQLEAWIHNTHSQPGARREARETSSRNALNVQAWRALPAPKIRRVQELCAGALDGLGRPYSEDEQRNALDLVLPGLNGFTWASSTASHPRN"

BASE COUNT 460 a 799 c 733 g 552 t

ORIGIN

Query Match 2.9% Score 59; DB 9; Length 2544;
Best Local Similarity 100.0%; Pred. No. 1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGCTCTACACGCCAGCTGTGCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
|||||
DB 1180 AGGCGTGGCGCTCTACACGCCAGCTGTGCTCAAGAGAGTGGCTTCTTCAACCTGCAG 1238

RESULT 13
AF246718 3278 bp mRNA linear PRI 31-OCT-2000
LOCUS Homo sapiens intestinal GlcNAc-6-sulfotransferase (CHST5) mRNA,
DEFINITION complete cds, alternatively spliced.
ACCESSION AF246718
VERSION AF246718.1 GI:11055254
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3278)
Akama,T.O., Nishida,K., Nakayama,J., Watanabe,H., Ozaki,K.,
Nakamura,T., Dote,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S.,
Fujiwara,T., Thonar,E.J., Shimomura,Y., Kinoshita,S., Tanigami,A.
and Fukuda,M.N.
Macular corneal dystrophy type I and type II are caused by distinct
mutations in a new sulphotransferase gene
Nat. Genet. 26 (2), 237-241 (2000)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 3278)
Akama,T.O., Nishida,K., Nakayama,J., Watanabe,H., Ozaki,K.,
Nakamura,T., Dote,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S.,
Fujiwara,T., Thonar,E.J., Shimomura,Y., Kinoshita,S.,
Tanigami,A. and Fukuda,M.N.
Direct Submission
Submitted (17-MAR-2000) Glycobiology Program, The Burnham
Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA

FEATURES
source
1..3278
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22: between D16S512 and D16S515"
1..3278
/gene="CHST5"
1..1396
/note="alternatively spliced"
1397..2632
/gene="CHST5"

CDS
5'UTR
gene
CDS

/note="I-GlcNAc6ST; carbohydrate sulfotransferase 5;
GlcNAc6ST expressed in small intestine and colon; similar
to N-acetylglucosamine 6-O-sulfotransferase encoded by
Genbank Accession Number AF176838; alternatively spliced"

BASE COUNT 686 a 988 c 866 g 738 t

ORIGIN

Query Match 2.9% Score 59; DB 9; Length 3278;
Best Local Similarity 100.0%; Pred. No. 1.1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGCTCTACACGCCAGCTGTGCTCAAGAGAGTGGCTTCTTCAACCTGCAG 736
|||||
DB 1950 AGGCGTGGCGCTCTACACGCCAGCTGTGCTCAAGAGAGTGGCTTCTTCAACCTGCAG 2008

RESULT 14
AF280086 3786 bp mRNA linear PRI 20-FEB-2001
LOCUS Homo sapiens N-acetylglucosamine 6-O-sulfotransferase GST-4beta
DEFINITION mRNA, complete cds.
ACCESSION AF280086
VERSION AF280086.1 GI:12060803
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS 1 (bases 1 to 3786)
Hemmerich,S., Lee,J.K., Bhakta,S., Bistrup,A., Ruddle,N.R. and
Kosen,S.D.
Chromosomal localization and genomic organization for the
galactose/N-acetylglactosamine/N-acetylglucosamine
6-O-sulfotransferase gene family
Glycobiology 11 (1), 75-87 (2001)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

2 (bases 1 to 3786)
Hemmerich,S., Bhakta,S., Lee,J.K., Bistrup,A., Ruddle,N.R. and
Kosen,S.D.
Direct Submission
Submitted (20-JUN-2000) Respiratory Diseases, Roche Bioscience,
3401 Hillview Ave., Palo Alto, CA 94304, USA

FEATURES
source
1..3786
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q23.1"
1..3786
/gene="GST4beta"
/note="clusters with the intestinal N-acetylglucosamine
6-O-sulfotransferase (GST4alpha) gene in Genbank Accession
Numbers AF176838 and AF176839"
189..1376
/gene="GST4beta"
/note="similar to N-acetylglucosamine 6-O-sulfotransferase
GST-4alpha"
/codon_start=1
/product="N-acetylglucosamine 6-O-sulfotransferase
GST-4beta"

CDS
gene
CDS


```

/protein_id="AAG48244.1"
/db_xref="GI:12060804"
/translation="MWLPRVSSITAVTALLAQTFLILIVSRPSSPGAGARVHL
VLSMRSGSSFGVGLFNQHPDVFLMEPMWHVTLISQSAATLHAAYDILVRVFLC
DMVDPAFLPMRRNLSDLFQMAVSRALCSPPACSAFPRASISSEAVCKPLCARQSTL
ARACRSYSHVLYKEVRFNLQVLYPLSDPALNRFIVLIVDPRAVLRSQOTAKAL
ARDNGIVLGNNGVWADPGLRVREVCRSYHVAIAAPLTKPPFLRGYRIYREEDL
ARPELAIKALVAFGLSLPLEAMVHNTTSGSGGARREAFKTSRRNALVDSQAMR
HAPFAKIRVQELCAGALQLCLGRVYSEDDQRNLADLVLPRLNGRTMASSTRSH
PRN"
polya_signal
3106..3111
/gene="GST4delta"
/note="alternative signal"
polya_signal
3734..3739
/gene="GST4beta"
BASE COUNT      848 a      1030 c      1104 g      804 t
ORIGIN
Query Match      2.9%; Score 59; DB 9; Length 3786;
Best Local Similarity 100.0%; Pred. No. 1,1e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CY      678 AGGCTGCGCCGTCTCAAGCCAGCGTGTCTCAAGAGGTCGCCCTCTTCAACCTGCAG 736
DB      676 AGGCTGCGCCGTCTCAAGCCAGCGTGTCTCAAGAGGTCGCCCTCTTCAACCTGCAG 734

RESULT 15
AF219991      71503 bp      DNA      linear      PRI 26-OCT-2000
LOCUS      AF219991
DEFINITION      Homo sapiens intestinal N-acetylglucosamine 6-O-sulfotransferase
                  (CHST5) and corneal N-acetylglucosamine 6-O-sulfotransferase
                  (CHST6) genes, complete cds.
ACCESSION      AF219991
VERSION        AF219991
KEYWORDS
SOURCE
ORGANISM      Homo sapiens.
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS      Akama,T.O., Nishida,K., Nakayama,T., Watanabe,H., Fujiwara,T.,
                Nakamura,T., Doi,A., Kawasaki,S., Inoue,Y., Maeda,N., Yamamoto,S.,
                Ozaki,K., Kinoshita,S., Shimomura,Y., Tanigami,A. and Fukuda,M.N.
                Macular corneal dystrophy type I and type II are caused by distinct
                mutations in a new sulfotransferase gene
JOURNAL      Nat. Genet. 26 (2), 237-241 (2000)
MEDLINE      20472330
PUBMED      11017086
REFERENCE
AUTHORS      Akama,T.O. and Fukuda,M.N.
TITLE      Direct Submission
JOURNAL      Submitted (29-DEC-1999) Glycobiology Program, The Burnham
                Institute, 10901 North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES
source
1..71503
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/map="16q22"
/clone="CITB-483K2"
358..461
/repeat_region
/repeat_family="L2"
/rpt_type-dispersed
complement(462..771)
/repeat_family="Alu"
/rpt_type-dispersed
772..945
/repeat_family="L2"
/rpt_type-dispersed
1169..1138
/repeat_family="MIR"
/rpt_type-dispersed
repeat_region
/rpt_type-dispersed

```

```

repeat_region      complement(1553..1804)
                    /rpt_family="Alu"
repeat_region      /rpt_type-dispersed
                    2616..2916
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    3309..3532
                    /note="similar to EST H12297 (EST cluster Hs.31147)"
                    3865..4028
                    /note="similar to EST H12297 (EST cluster Hs.31147)"
                    4099..4406
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    4793..5250
                    /rpt_family="L2"
                    /rpt_type-dispersed
                    complement(5251..5329)
                    /rpt_family="MIR"
                    /rpt_type-dispersed
                    complement(5768..6080)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    complement(6081..6265)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    6111..6567
                    /note="similar to EST AA001593 (EST cluster Hs.6853)"
                    8035..8350
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    10345..11049
                    /note="similar to EST A1950923 (EST cluster Hs.6853)"
                    11043..11195
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    11200..11222
                    /rpt_family="simple_repeat"
                    /rpt_type-tandem
                    11570..11671
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    complement(12065..12198)
                    /rpt_family="L1"
                    /rpt_type-dispersed
                    complement(12205..12493)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    complement(12813..13110)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    14185..20821
                    /gene="CHST5"
                    join(14185..14228,16169..16263,17712..20821)
                    /gene="CHST5"
                    /product="intestinal N-acetylglucosamine
                    6-O-sulfotransferase"
                    14681..14713
                    /rpt_family="simple_repeat"
                    /rpt_type-tandem
                    complement(14715..14990)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    complement(14995..15288)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    complement(15292..15589)
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    15668..15718
                    /rpt_family="Alu"
                    /rpt_type-dispersed
                    15977..16098
                    /rpt_family="MIR"

```



```

repeat_region /rpt_type=dispersed
complement(17298..17433)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region complement(18001..18310)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(18420..18591)
/rpt_family="MIR"
/rpt_type=dispersed
repeat_region 18732..18767
/rpt_family="Simple-repeat"
/rpt_type=tandem
18967..20202
/gene="CHST5"
/note="I-GlcNAc6ST; carbohydrate sulfoltransferase 5;
expressed in the small intestine and colon"
/codon_start=1
/product="intestinal
N-acetylglucosamine-6-O-sulfoltransferase"
/db_xref="GI:11023148"
/translation="MGMRARVRYVAHSRRPPAARMMLPRFSSTKVTVLLAOTCIL
LFTSRPSSSPAGDEPRVHYLVLSRPSSTGSCLESQHPDYFLMEPRHWITL
SOGSNATIHMAVRDLMRSTFLCDMDYFAIMPQSRNLSAEFNWATSRALCSPPACSAF
PRGTSKQDVCKTCTQRPVSLAREACSYSHVVKERFENLQVLPPLSDPLNR
IVHLVRPRAVLRSRSEAGPILANDNGIVLTNGKWEADPHRLIRVCRSHYRIE
AATLKPPFLRGRLVRFEDLAREPLAELIYALYGLITLPDLEAMINHTGSGI
GKPIEAFSTSRNARNVQAWRHAPFKILRVQEVCAQLLLGYRVYSADQQRDL
TLDLVPRGPHFMSAFSD"
20504..20806
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 20869..20961
/rpt_family="L1"
/rpt_type=dispersed
repeat_region 20962..21200
/rpt_family="MER4-group"
/rpt_type=dispersed
21325..21429
/rpt_family="MER4-group"
/rpt_type=dispersed
repeat_region 21430..21722
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 21725..22007
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 22008..22040
/rpt_family="Simple-repeat"
/rpt_type=tandem
22041..22313
/rpt_family="MER4-group"
/rpt_type=dispersed
repeat_region 22325..22593
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region 22594..22949
/rpt_family="MER4-group"
/rpt_type=dispersed
repeat_region 23116..23221
/rpt_family="Simple-repeat"
/rpt_type=tandem
23344..23400
/rpt_family="MALR"
/rpt_type=dispersed
repeat_region 23433..23455
/rpt_family="Simple-repeat"
/rpt_type=tandem
repeat_region complement(23457..23720)
/rpt_family="Alu"
/rpt_type=dispersed
repeat_region complement(23789..24140)

```

```

/rpt_family="MER21-group"
/rpt_type=dispersed
24141..24287
/rpt_family="MALR"
/rpt_type=dispersed
repeat_region complement(24308..24420)
/rpt_family="MALR"
/rpt_type=dispersed
repeat_region complement(24467..24528)
/rpt_family="MER1_type"

Query Match      2.9%; Score 59; DB 9; Length 71503;
Best Local Similarity 100.0%; Pred. No. 1.2e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 678 AGCGCTGCGGCTCCACACGACGCGTGCACAGAGAGTGGCTTCTCAACCTGCAG 736
Db 19520 AGCGCTGCGGCTCCACACGACGCGTGCACAGAGAGTGGCTTCTCAACCTGCAG 19578

RESULT 16
AC009163/c 157337 bp DNA linear PRI 02-NOV-2001
LOCUS AC009163
DEFINITION Homo sapiens chromosome 16 clone RP11-77K12, complete sequence.
ACCESSION AC009163
VERSION AC009163.5 GI:16596526
KEYWORDS HTG.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 157337)
DOE Joint Genome Institute.
Sequencing Of Human Chromosome 16
Unpublished
2 (bases 1 to 157337)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 157337)
DOE Joint Genome Institute.
Direct Submission
Submitted (02-NOV-2001) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Nov 2, 2001 this sequence version replaced gi:17689976.
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
Sequence Quality Assessment:
This entry has been annotated with sequence quality
estimates computed by the Phrap assembly program.
All manually edited bases have been reduced to quality zero.
Quality levels above 40 are expected to have less than
1 error in 10,000 bp.
Base-by-base quality values are not generally visible from the
Genbank flat file format but are available as part
of this entry's ASN.1 file.
-----
FEATURES
Source Location/Qualifiers
1..157337
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-77K12"

BASE COUNT 45295 a 36632 c 36308 g 39102 t

```

ORIGIN

Query Match 2.9%; Score 59; DB 9; Length 157337;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

678 AGGCGTCCGCTCTACAGCAGTGTGCTCAAGAGGAGCGCTTCTTCAACCTGCAG 736
|||||
DB 144095 AGGCGTCCGCTCTACAGCAGTGTGCTCAAGAGGAGGAGCGCTTCTTCAACCTGCAG 144037

RESULT 17
AC025287 194832 bp DNA 1linear PRI 26-JAN-2002
DEFINITION Homo sapiens chromosome 16 clone RP11-490B18, complete sequence.
ACCESSION AC025287 GI:18376863
VERSION
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
JOURNAL Unpublished
2 (bases 1 to 194832)
DOE Joint Genome Institute.
AUTHORS
TITLE Direct Submission
REFERENCE
JOURNAL Submitted (08-MAR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
3 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
JOURNAL Submitted (21-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
4 (bases 1 to 194832)
DOE Joint Genome Institute and Stanford Human Genome Center.
REFERENCE
JOURNAL Direct Submission
Submitted (26-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
On Jan 26, 2002 this sequence version replaced gi:17976465.
COMMENT
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.4% of sequence;
Estimated Total Number of Errors is 0.6.
FEATURES
source
1.194832
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-490B18"
BASE COUNT 53784 a 43240 c 43815 g 53993 t
ORIGIN

Query Match 2.9%; Score 59; DB 9; Length 194832;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

678 AGGCGTCCGCTCTACAGCAGTGTGCTCAAGAGGAGCGCTTCTTCAACCTGCAG 736
|||||
DB 188038 AGGCGTCCGCTCTACAGCAGTGTGCTCAAGAGGAGGAGCGCTTCTTCAACCTGCAG 188096

RESULT 18
AC009105 208185 bp DNA 1linear HTG 25-APR-2001
LOCUS Homo sapiens chromosome 16 clone RP11-455E15, WORKING DRAFT
DEFINITION
SEQUENCE 17 unordered pieces.
ACCESSION AC009105
VERSION AC009105.7 GI:13786304

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ACTIVE1N.
Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 208185)
DOE Joint Genome Institute
Sequencing of Human Chromosome 16
Unpublished
2 (bases 1 to 208185)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Apr 25, 2001 this sequence version replaced gi:7689928.
-----Genome Center
Center: Joint Genome Institute
Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 0
Center clone name: RPCT-11-455E15

Summary Statistics
Consensus quality: 180040 bases at least Q40
Consensus quality: 192095 bases at least Q30
Consensus quality: 195726 bases at least Q20
Estimated insert size: 195850; agarose-gel estimation
Estimated insert size: 206585; sum-of-contigs estimation
Quality coverage: 6.57 in Q20 bases; agarose-gel estimation
Quality coverage: 6.23 in Q20 bases; sum-of-contigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1062: contig of 1062 bp in length
* 1063 1162: gap of unknown length
* 1163 2359: contig of 1197 bp in length
* 2359 2459: gap of unknown length
* 2459 2600: contig of 1194 bp in length
* 2600 3653: gap of unknown length
* 3653 3753: gap of unknown length
* 3753 3754: gap of 1887 bp in length
* 3754 5640: contig of unknown length
* 5640 5740: gap of unknown length
* 5740 5741: contig of 1630 bp in length
* 5741 7371: gap of unknown length
* 7371 7471: gap of 1114 bp in length
* 7471 8584: contig of unknown length
* 8584 8684: gap of unknown length
* 8684 10014: contig of 1330 bp in length
* 10014 10115: gap of unknown length
* 10115 12451: contig of 2337 bp in length
* 12451 12551: gap of unknown length
* 12551 28715: contig of 16164 bp in length
* 28715 28716: gap of unknown length
* 28716 28815: gap of unknown length
* 28815 43340: contig of 14325 bp in length
* 43340 43241: gap of unknown length
* 43241 53865: contig of 10625 bp in length
* 53865 53965: gap of unknown length
* 53965 72091: contig of 18126 bp in length
* 72091 72192: gap of unknown length
* 72192 87356: contig of 15164 bp in length
* 87356 87357: gap of unknown length
* 87357 87456: gap of unknown length
* 87456 109461: contig of 22006 bp in length
* 109461 109561: gap of unknown length
* 109561 131379: contig of 21818 bp in length
* 131379 131479: gap of unknown length
* 131479 131480: contig of 26077 bp in length
* 131480 157557: contig of unknown length
* 157557 157558: gap of unknown length
* 157558 208185: contig of 50529 bp in length.

FEATURES
source Location/Qualifiers
1..208185
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="16"
/clone="RP11-455E15"
/clone_id="Rpci human BAC library 11"
BASE COUNT 58183 a 45373 c 45592 g 57204 t 1833 others
ORIGIN

Query Match 2.9%; Score 59; DB 2; Length 208185;
Best Local Similarity 100.0%; Pred. No. 1.3e-22;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 678 AGGCGTCCGCTCTCTACAGCCAGTGGTCTCAGAGAGTGGCTTTCTCACTGCAG 736
|||||
Db 107982 AGGCGTCCGCTCTCTACAGCCAGTGGTCTCAGAGAGTGGCTTTCTCACTGCAG 108040
|||||

RESULT 19
AR203340 37 bp DNA linear PAT 20-JUN-2002
LOCUS
DEFINITION Sequence 7 from patent US 6365365.
ACCESSION AR203340
VERSION AR203340.1 GI:21499704
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 37)
AUTHORS Bistrup, A., Rosen, S.D., Tangemann, K. and Hemmerich, S.
TITLE Method of determining whether an agent modulates glycosyl
JOURNAL Patent: US 6365365-A 7 02-APR-2002;
FEATURES Location/Qualifiers
source 1..37
BASE COUNT 14 a 9 c 8 g 6 t
ORIGIN

Query Match 1.8%; Score 37; DB 6; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 884 AAACCTAAGAAGAGAGACCAACCTACTGTGATGC 920
|||||
Db 1 AAACCTAAGAAGAGAGACCAACCTACTGTGATGC 37
|||||

RESULT 20
AF131236 1926 bp DNA linear ROD 10-AUG-1999
LOCUS
DEFINITION Mus musculus N-acetylglucosamine 6-O-sulfotransferase gene,
ACCESSION AF131236
VERSION AF131236.1 GI:4927115
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1926)
AUTHORS Bistrup, A., Bhakta, S., Lee, J.K., Belov, Y.Y., Gunn, M.D., Zuo, F.R.,
Huang, C.C., Kannagi, R., Rosen, S.D. and Hemmerich, S.
TITLE Sulfotransferases of two specificities function in the
reconstitution of high endothelial cell ligands for L-selectin
JOURNAL J. Cell Biol. 145 (4), 899-910 (1999)
MEDLINE 99264336
PUBMED 10330415
REFERENCE 2 (bases 1 to 1926)
AUTHORS Bistrup, A., Tangemann, K., Bhakta, S., Lee, J.-K., Belov, Y.Y.,
Gunn, M.D., Zuo, F.-R., Huang, C.-C., Kannagi, R., Rosen, S.D. and
Hemmerich, S.

TITLE Direct Submission
JOURNAL Submitted (25-FEB-1999) Respiratory Diseases, Roche Bioscience,
3401 Hillview Avenue, Palo Alto, CA 94304, USA
FEATURES
source Location/Qualifiers
1..1926
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/cell_type="high endothelial"
/tissue_type="tonsil"
<405..>1571
/product="N-acetylglucosamine 6-O-sulfotransferase"
405..1571
/function="biosynthesis of endothelial ligands for
L-selectin"
/note="HEC-G1CnAc-6-ST"
/codon_start=1
/product="N-acetylglucosamine 6-O-sulfotransferase"
/protein_id="P4033016.1"
/db_xref="GI:4927116"
/translation="MMLKKRLLMEFGSQVIVVAFIMSVSRHLSORESRPVAV
LVLSMRSGSPVGOLEFGHPDVEFLKEPAMVCFSTAWKLRHVAVDLRSVEL
CDMSVPDAPVNNPGPRKQSSILFQWROSALCSAIVCPFPAPHTLSPRKCKLGGQDF
DMEKACRSRGFPVYLVKREVRFLSDALYPLITDPSLNIHVHLYRDPRAVFRSEHTI
ELWDSHIVIGQHEJTEKEDQPIAKTIKCSQVDIVKAIQTLPELQOQRYFLFVE
DLVAPLAQTTFLRYKFGVLDLPHLQTVYVNVYRGKMGQGHAFINRNALNYSQMR
WSLPYEVSOLODPCGEMDLGLGYOVRSQOEOGNLSLDLSSSHLIGOVFREG"

BASE COUNT 426 a 520 c 484 g 496 t
ORIGIN

Query Match 1.3%; Score 27; DB 10; Length 1926;
Best Local Similarity 100.0%; Pred. No. 0.0013;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 494 GACATGAGCGTCTTTGATGCTACATG 520
|||||
Db 714 GACATGAGCGTCTTTGATGCTACATG 740
|||||

RESULT 21
AF109155 2201 bp mRNA linear ROD 27-JUL-1999
LOCUS
DEFINITION Mus musculus L-selectin ligand sulfotransferase mRNA, complete cds.
ACCESSION AF109155
VERSION AF109155.1 GI:5596405
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2201)
AUTHORS Hiraoka, N., Petryniak, B., Nakayama, J., Tsudoi, S., Suzuki, M.,
Yeh, J.C., Izawa, D., Tanaka, T., Miyasaka, M., Lowe, J.B. and Fukuda, M.
TITLE A novel, high endothelial venule-specific sulfotransferase
expresses 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed
by CD34
JOURNAL Immunity 11 (1), 79-89 (1999)
MEDLINE 99361934
PUBMED 10435581
REFERENCE 2 (bases 1 to 2201)
AUTHORS Hiraoka, N. and Fukuda, M.
TITLE Direct Submission
JOURNAL Submitted (23-NOV-1998) Glycobiology, The Burnham Institute, 10901
North Torrey Pines Road, La Jolla, CA 92037, USA
FEATURES
source Location/Qualifiers
1..2201
/organism="Mus musculus"
/db_xref="taxon:10090"
319..1485
/function="directs expression of 6-sulfo sialyl Lewis x,
an L-selectin ligand displayed by CD34"
/note="LSSR"
/codon_start=1

```

/product="L-selectin ligand sulfotransferase"
/protein_id="A045579.1"
/db_xref="GI:5396406"
/translation="MMLKKGRLNFGSQVIYVALFIMSVHRHLSQREESRRPYVH
LVLSWRSGSSSEYVGLFGQHPDYFLMEPAWHYMTFTSTAKLHMAYRDLRSYFL
CDMSVFDAYNNPGPRKSSLPQWESRALCSAPVCDFFPAHEISSPKHCKLLCGOOPF
DVEKACRSHGEVFLKEVRPLSQALYPLTDPSLNLHYVHLVYRDPRAVRSREHHTI
ELVVDSHIVIGCHLFTIKREDOPYAMKTIICKSNDIVKAIOTLPEALQORYLFLAYE
DIYRAFLAQTTRLIKTFVGLDFLPHLOTWYHNTTRGKMGQHAFTIYARNALNYSQMR
WSLPYERVSQLDAGCEAMDLLGYLOVRSQDEQGNLSDLLSSSHLGOVREG"

```

```

BASE COUNT      537 a      629 c      534 g      501 t
ORIGIN
Query Match      1.3%: Score 27; DB 10; Length 2201;
Best Local Similarity 100.0%: Pred. No. 0.0015;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 494 GACATGAGCGCTTTGATGCTACATG 520
      |||
Db 628 GACATGAGCGCTTTGATGCTACATG 654

```

Search completed: February 22, 2003, 22:24:25
 Job time : 6247 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 18:56:30 ; Search time 444 Seconds

(without alignments)
10306.443 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032 1 ggcctcgagccagagatgcct.....ataagaaaaaaaaaaaaa 2032

Sequence: 1 ggcctcgagccagagatgcct.....ataagaaaaaaaaaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Capext 60.0

Searched: 2185239 seqs, 112599159 residues

Word size : 25

Total number of hits satisfying chosen parameters: 21

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2032	100.0	2032	20	AA220792	Human glycosyl sul
2	1802	88.7	1979	22	AAK94229	Human full-length
3	1615	79.5	2065	21	AA294211	Human transferrase
4	1211	59.6	1333	24	AA516947	Human L-selectin s
5	741	36.5	877	22	AAK91803	Human CDNA 5'-end
6	741	36.5	877	22	AAK93921	Human CDNA clone r
7	505	24.9	517	24	ABK54724	Human colon cancer
8	351	17.3	548	22	AAK92568	Human CDNA 3'-end
9	306	15.1	2988	21	AACT6156	Human OREF1711

10	286	14.1	389	24	ABK54794	Human colon Cancer
11	60	3.0	60	24	ABN38074	Human spliced tran
12	59	2.9	1647	24	AAAD24670	Human drug metabol
13	59	2.9	1694	22	AAAD02700	Human glycosyl sul
14	59	2.9	2044	22	AAAD02699	Human glycosyl sul
15	59	2.9	2170	22	AAAD02698	Human glycosyl sul
16	59	2.9	2544	24	ABN89506	Human glycosyl sul
17	59	2.9	4836	24	ABN89533	Human corneal N-ac
18	59	2.9	160552	22	AAAD02697	Human glycosyl sul
19	37	1.8	37	20	AAZ20798	PCR primer for gly
20	30	1.5	30	24	ABK70821	Probe for human ge
21	27	1.3	1926	20	AAZ20793	Mouse glycosyl sul

ALIGNMENTS

RESULT 1
AAZ20792
ID AAZ20792 standard, DNA; 2032 BP.
XX
AC AAZ20792;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human glycosyl sulfotransferase-3 coding sequence.
XX
KW glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KW selectin binding interaction; inflammation; lymphocyte homing; human;
KW secondary lymph organ; ss.
XX
OS Homo sapiens.
XX
PN WO949018-A1.
XX
PD 30-SEP-1999.
XX
PF 26-FEB-1999; 95WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIV CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistup A, Rosen SD, Tangemann K, Hemmerich S;
XX WPI, 1999-580442/49.
DR P-FSDB; AAY39918.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
XX polynucleotides -
XX
PS Claim 4; Fig 1; 59pp; English.
XX
XX This sequence encodes the human glycosyl sulfotransferase-3 (GST-3) of
CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC The sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions;
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.
XX
SQ Sequence 2032 BP; 468 A; 569 C; 490 G; 505 T; 0 other;
XX
Query Match 100.0%; Score 2032; DB 20; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGCCTCGAGCCAGAGATGCCTCCAGTCGGGGAATAATGCTTCATTTGCTCCAG 60
DB 1 GGCCTCGAGCCAGAGATGCCTCCAGTCGGGGAATAATGCTTCATTTGCTCCAG 60

QY 61 CCCACCTCAGACAGTCTCCACCACCTTGAGTCTCAGACAGTGTAAAGCTGTACTTTCA 120
DB 61 CCACACCTCAGACAGTCTCCACCACCTTGAGTCTCAGACAGTGTAAAGCTGTACTTTCA 120
QY 121 CACCTTCCTGGAGCGAGTGTCTTCTCAAGCCGCTTGCAGAGTCTTCCACTTACGAC 180
DB 121 CACCTTCCTGGAGCGAGTGTCTTCTCAAGCCGCTTGCAGAGTCTTCCACTTACGAC 180
QY 181 AATGCTACTGCTTAAATAATGAAAGCTGCTGTCTGTCTGTTTCCAGATGGCATCTT 240
DB 181 AATGCTACTGCTTAAATAATGAAAGCTGCTGTCTGTCTGTTTCCAGATGGCATCTT 240
QY 241 GGGCTATCTTCTCAGATGTACAGCCCAACATCAGCTCCCTGTCTATGAAAGCAGCC 300
DB 241 GGGCTATCTTCTCAGATGTACAGCCCAACATCAGCTCCCTGTCTATGAAAGCAGCC 300
QY 301 CGAGCCATGACAGTGTCTGTTCTGTCTCAGAGCCCTGAGCTCTCTTCTTGTGGGCA 360
DB 301 CGAGCCATGACAGTGTCTGTTCTGTCTCAGAGCCCTGAGCTCTCTTCTTGTGGGCA 360
QY 361 GCTTTTGGGCAACCCAGATGTCTTCTACCTGATGAGCCCGCTGCGACGTGTGAT 420
DB 361 GCTTTTGGGCAACCCAGATGTCTTCTACCTGATGAGCCCGCTGCGACGTGTGAT 420
QY 421 GACCTTCAAGCAGACAGCCGCTGATGCTGACATGGCTGTGGGATCTGATACGGGC 480
DB 421 GACCTTCAAGCAGACAGCCGCTGATGCTGACATGGCTGTGGGATCTGATACGGGC 480
QY 481 CGTCTTCTGTGGACATGAGCGTCTTGTATGCTACATGGAACCTGTCTCCCGAGACA 540
DB 481 CGTCTTCTGTGGACATGAGCGTCTTGTATGCTACATGGAACCTGTCTCCCGAGACA 540
QY 541 GTCCAGCTCTTCACTGGGAGAAAGCCGGGCGCTGTCTGTGACCTGCTGTGACAT 600
DB 541 GTCCAGCTCTTCACTGGGAGAAAGCCGGGCGCTGTCTGTGACCTGCTGTGACAT 600
QY 601 CATCCCAAGATGAATCATCCCCGGGCTCAGTCAAGCTCTGTGACATCAACAGCC 660
DB 601 CATCCCAAGATGAATCATCCCCGGGCTCAGTCAAGCTCTGTGACATCAACAGCC 660
QY 661 CTTTGAAGTGTGGAGAAAGCCCTGCTCTACAGCCAGTGTCTTAAAGAGTGGC 720
DB 661 CTTTGAAGTGTGGAGAAAGCCCTGCTCTACAGCCAGTGTCTTAAAGAGTGGC 720
QY 721 CTTCTTCAACCTGACATGCTCTACCCGCTGTGAAGAACCCTCTCAACCTGATAT 780
DB 721 CTTCTTCAACCTGACATGCTCTACCCGCTGTGAAGAACCCTCTCAACCTGATAT 780
QY 781 CGTGACCTGTGTCCGGAGCCCGGGCGTGTCCGTCGCGAGAACGCAAAAGGAGA 840
DB 781 CGTGACCTGTGTCCGGAGCCCGGGCGTGTCCGTCGCGAGAACGCAAAAGGAGA 840
QY 841 TGTCAATGATGACAGTGTGTGATGGGAGAGCATGACAGAAACTCAAGAGAGGA 900
DB 841 TGTCAATGATGACAGTGTGTGATGGGAGAGCATGACAGAAACTCAAGAGAGGA 900
QY 901 CCAACCTCTATGTGATGATGATCTGCAAAAGCAGTGTGAGATCTACAAAGCAT 960
DB 901 CCAACCTCTATGTGATGATGATCTGCAAAAGCAGTGTGAGATCTACAAAGCAT 960
QY 961 CCAAGTCTTGGCCCAAGGCGCTGAGAGAGCAAGCTGTGGCTATGAGAGCTGGC 1020
DB 961 CCAAGTCTTGGCCCAAGGCGCTGAGAGAGCAAGCTGTGGCTATGAGAGCTGGC 1020
QY 1021 TGGAGCCCTGTGGCCCAAGCTTCCGAAATGTATGATTCGTGGGATTTGATTCGC 1080
DB 1021 TGGAGCCCTGTGGCCCAAGCTTCCGAAATGTATGATTCGTGGGATTTGATTCGC 1080
QY 1081 CCAATCTTCAAGCTGTGGTGTATATCAATCACCCGAGGCAAGGCGATGGGTGACAGCTTT 1140
DB 1081 CCAATCTTCAAGCTGTGGTGTATATCAATCACCCGAGGCAAGGCGATGGGTGACAGCTTT 1140

QY 1141 CCACAAATGCGCAGGAGTGGCCTTAATGTCTCCAGGCTTGGCGTGTCTTGGCCTTA 1200
DB 1141 CCACAAATGCGCAGGAGTGGCCTTAATGTCTCCAGGCTTGGCGTGTCTTGGCCTTA 1200
QY 1201 TGAAGAAGTGTCTGCACTTCAAGAAAGCCTGTGGGATGTCATGAAATTTGCTGGGCTACCG 1260
DB 1201 TGAAGAAGTGTCTGCACTTCAAGAAAGCCTGTGGGATGTCATGAAATTTGCTGGGCTACCG 1260
QY 1261 CCAGCTCAGATCTGAAACAGAAAGAGAAACCTGTGTGATGCTGTCTGTACTGAGAC 1320
DB 1261 CCAGCTCAGATCTGAAACAGAAAGAGAAACCTGTGTGATGCTGTCTGTACTGAGAC 1320
QY 1321 TGTCCCTAGCAAAATCCACTAAGAGGTTGAGAAAGCTTCTCTCCACCTGTGTGAGCC 1380
DB 1321 TGTCCCTAGCAAAATCCACTAAGAGGTTGAGAAAGCTTCTCTCCACCTGTGTGAGCC 1380
QY 1381 TCAGTCACTTCTGTGATGCTGTGAGCCTTGGCTACATCTGAGCCTTAATACATG 1440
DB 1381 TCAGTCACTTCTGTGATGCTGTGAGCCTTGGCTACATCTGAGCCTTAATACATG 1440
QY 1441 TCTGTGGGTATCACTAGTGTGAGTGTGTCTCCACAGCTGTCTAAGCAGAGACTTTT 1500
DB 1441 TCTGTGGGTATCACTAGTGTGAGTGTGTCTCCACAGCTGTCTAAGCAGAGACTTTT 1500
QY 1501 GTGTCCATGCTGTGTCTAGAAACAGACTGGGAACTTATGTGAGAGCAGATCCAC 1560
DB 1501 GTGTCCATGCTGTGTCTAGAAACAGACTGGGAACTTATGTGAGAGCAGATCCAC 1560
QY 1561 CAGTGAACAGGGTATGCTCTCTCTTCTTCTGTGATCTCTGTGTGGCAGACTTCA 1620
DB 1561 CAGTGAACAGGGTATGCTCTCTCTCTTCTTCTGTGATCTCTGTGTGGCAGACTTCA 1620
QY 1621 AGACTTGTGGCCTGAGAGGCTTATTAAGCAGACAGATCAGTGAATGTATGATCAATA 1680
DB 1621 AGACTTGTGGCCTGAGAGGCTTATTAAGCAGACAGATCAGTGAATGTATGATCAATA 1680
QY 1681 AACTTCCCTGTCCACATCTTGGCCCAATGGGATGATGATCTTCAACCAAGAGCTCACAGC 1740
DB 1681 AACTTCCCTGTCCACATCTTGGCCCAATGGGATGATGATCTTCAACCAAGAGCTCACAGC 1740
QY 1741 ATTTTCCACAGATGCAATTTCTGAGCCCTTGGAGTTCCAGTGTGATTCAGAGAAAGAA 1800
DB 1741 ATTTTCCACAGATGCAATTTCTGAGCCCTTGGAGCTTCCAGTGTGATTCAGAGAAAGAA 1800
QY 1801 GTGGAAACAAAGTGTGATGCTTACTTATGACCTTGAACATCAGAGCTATGATGATA 1860
DB 1801 GTGGAAACAAAGTGTGATGCTTACTTATGACCTTGAACATCAGAGCTATGATGATA 1860
QY 1861 AATATGAACAAATCTGTGCAAAAGAGCAAGCTTTAAGTTCAAGGCTGCTGGGC 1920
DB 1861 AATATGAACAAATCTGTGCAAAAGAGCAAGCTTTAAGTTCAAGGCTGCTGGGC 1920
QY 1921 TGCATTTGAATATCACTTCCCTCTGCAATTTCCATCACAATAAGACTTGAAGCTGTG 1980
DB 1921 TGCATTTGAATATCACTTCCCTCTGCAATTTCCATCACAATAAGACTTGAAGCTGTG 1980
QY 1981 AAGCTGCATCTGTATATCTAATAATCCAAATGAAGAAAAA 2032
DB 1981 AAGCTGCATCTGTATATCTAATAATCCAAATGAAGAAAAA 2032
RESULT 2
AAK94229
ID AAK94229 standard; cDNA; 1979 BP.
AC AAK94229;
XX 06-NOV-2001 (first entry)
XX Human full-length cDNA, SEQ ID NO: 2815.
DE Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
XX

OS Homo sapiens.
XX
XX EPI130094-A2.
XX
XX 05-SEP-2001.
XX
XX 07-JUL-2000; 2000EP-0114089.
XX
XX 08-JUL-1999; 99JP-0194486.
XX 11-JAN-2000; 2000JP-018774.
XX 02-MAY-2000; 2000JP-0183765.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX WPI: 2001-524255/58.
XX P-PSDB: AAM93309.
XX
XX 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
XX
XX Claim 8; SEQ ID NO 2816; 1380bp + sequence listing; English.
XX
XX The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' - and 3' -ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a full length
CC human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX
XX Sequence 1979 BP; 451 A; 542 C; 474 G; 512 T; 0 other;
SQ
Query Match 88.7%; Score 1802; DB 22; Length 1979;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1902; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 114 ACTTTCACAGCTCTCTGGAGCAGAGCTTCTCAAGCCGCTTTCAGAGCTTTCACCT 173
Db 1 ACTTTCACAGCTCTCTGGAGCAGAGCTTCTCAAGCCGCTTTCAGAGCTTTCACCT 60
QY 174 TCAGACACATGCTACTGCTAAATAATGAAAGTCTGCTGTTTCTGGTTCCAGATGG 233
Db 61 TCAGACACATGCTACTGCTAAATAATGAAAGTCTGCTGTTTCTGGTTCCAGATGG 120
QY 234 CCATCTGGAGCTCTATCTTCTTCACATGTATACAGCACAACATAGTCCCTGCTATGAGG 293
Db 121 CCACTCTGGAGCTCTATCTTCTTCACATGTATACAGCACAACATAGTCCCTGCTATGAGG 180
QY 294 CACAGCCCGAGCGATGACGTGCTGTTCTGCTTCTCTGCGCTGCTGCTCTCTTTTG 353
Db 181 CACAGCCCGAGCGATGACGTGCTGTTCTGCTTCTCTGCGCTGCTGCTCTCTTTTG 240
QY 354 TGGGGAGAGCTTTTGGGGAGAGCACCAGATGTTTCTACCTATATGAGCCCGCTGGACG 413
Db 241 TGGGGAGAGCTTTTGGGGAGAGCACCAGATGTTTCTACCTATATGAGCCCGCTGGACG 300
QY 414 TGTGATGACCTTCAGACAGACACCGCTGATGCTGACATAGCTGCTGGGATCTGA 473
Db 301 TGTGATGACCTTCAGACAGACACCGCTGATGCTGACATAGCTGCTGGGATCTGA 360
QY 474 TACGGGGCGCTTCTTGTGTGGGACATGAGCGCTTTGATGCTTAATGGAACCTGCTCCC 533
Db 361 TACGGGGCGCTTCTTGTGTGGGACATGAGCGCTTTGATGCTTAATGGAACCTGCTCCC 420
QY 534 GGAGACAGTCCAGCCTTTTCATGTGGGAGAACGCCGGGCGCTGTGTTCTGACCTGCT 593

Db 421 GGAGACAGTCCAGCCTTTTCATGTGGGAGAACGCCGGGCGCTGTGTTCTGACCTGCT 480
QY 594 GTGACATCATCCCAAGAGATGAAATCATCCCCGGGGCTCACTGAGGCTCTTGTGACATC 653
Db 481 GTGACATCATCCCAAGAGATGAAATCATCCCCGGGGCTCACTGAGGCTCTTGTGACATC 540
QY 654 AACAGCCCTTTGAGAGTGTGGAGAGAGGCTGCGCTCTCTACAGCAGAGTGTGCTCAAG 713
Db 541 AACAGCCCTTTGAGAGTGTGGAGAGAGGCTGCGCTCTCTACAGCAGAGTGTGCTCAAG 600
QY 714 AGGTGGCTTCTTCACTGACCTGACGTCCCTCTACCCGCTGTGAAGAACCCCTCCCTAAC 773
Db 601 AGGTGGCTTCTTCACTGACCTGACGTCCCTCTACCCGCTGTGAAGAACCCCTCCCTAAC 660
QY 774 TGCATATCGTGCACCTGCTGCGGAGCCCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 833
Db 661 TGCATATCGTGCACCTGCTGCGGAGCCCGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 834 AGGAGATCTCATGATTGACAGTGCATTTGTATGGGGCAGATGAGCAGAACTCAGAG 893
Db 721 AGGAGATCTCATGATTGACAGTGCATTTGTATGGGGCAGATGAGCAGAACTCAGAG 780
QY 894 AGGAGACCAACCTTACTATGTATGATGAGTTCATCTGCCAAAGCCAGCTGAGATCTACA 953
Db 781 AGGAGACCAACCTTACTATGTATGATGAGTTCATCTGCCAAAGCCAGCTGAGATCTACA 840
QY 954 AGACATTCAGTCTCTTGGCCCAAGGCCCTGCGAGAACGCTACCTGCTGCTGCTATGAG 1013
Db 841 AGACATTCAGTCTCTTGGCCCAAGGCCCTGCGAGAACGCTACCTGCTGCTGCTATGAG 900
QY 1014 ACCGTGCTGAGCCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1073
Db 901 ACCGTGCTGAGCCCTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 1074 TCTTGCCCATCTTTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1133
Db 961 TCTTGCCCATCTTTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
QY 1134 ACCGTTTCCACACAAATGCCAGGAGTCCCTTAATGCTTCCAGGCTTGGCGTGTCTT 1193
Db 1021 ACCGTTTCCACACAAATGCCAGGAGTCCCTTAATGCTTCCAGGCTTGGCGTGTCTT 1080
QY 1194 TGCCCTATGAAAAGGTTTCTCGACTTCAGAAACCTGTGGCGATGCGATTAATTTGCTG 1253
Db 1081 TGCCCTATGAAAAGGTTTCTCGACTTCAGAAACCTGTGGCGATGCGATTAATTTGCTG 1140
QY 1254 GCTACGCGCAGTGCATGATGGAACAGAGAAACCTGTTGCTGATCTTCTGCTTA 1313
Db 1141 GCTACGCGCAGTGCATGATGGAACAGAGAAACCTGTTGCTGATCTTCTGCTTA 1200
QY 1314 CCTGAGCTGTCCTGAGCAATCCAGTAAGAGGTTGAGAAAGCTTTCGCTGCACTGGT 1373
Db 1201 CCTGAGCTGTCCTGAGCAATCCAGTAAGAGGTTGAGAAAGCTTTCGCTGCACTGGT 1260
QY 1374 GTAGGCTCAGTCACTTCTCTGAATGCTTCTAGGCTTCTCTACATCTCTGAGCTTAA 1433
Db 1261 GTAGGCTCAGTCACTTCTCTGAATGCTTCTAGGCTTCTCTACATCTCTGAGCTTAA 1320
QY 1434 CTACATGCTGTGGGTATCTACATGATGATGATGATGATGATGATGATGATGATGAT 1493
Db 1321 CTACATGCTGTGGGTATCTACATGATGATGATGATGATGATGATGATGATGATGAT 1380
QY 1494 GACTTTTGTGTCATGCTTCTGTAGAAAACAGACCTGGGAGACCTTAATGATGAGACAG 1553
Db 1381 GACTTTTGTGTCATGCTTCTGTAGAAAACAGACCTGGGAGACCTTAATGATGAGACAG 1440
QY 1554 ATCCACACATGAAAGAGGATATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1613
Db 1441 ATCCACACATGAAAGAGGATATGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1500
QY 1614 ACTTCAGAGACTTGTGGCTGAGAGGCTTATTAAGACAGACAGATGATGATGATGATGAT 1673

	<p>The present sequence is that of cDNA clone 2617407C1 encoding human transferase TRNSF-11 (see AAY/92219), 1 of 15 claimed human transferase proteins of the invention (see AAY/9209-23). The clone was isolated from gall bladder cDNA library GBLANOT01. TRNSF-11 is expressed in dermalogic and gastrointestinal tissues, especially those associated with inflammation and cell proliferation. It shows homology to mouse N-acetylglucosamine 6-O-sulfotransferase. A polynucleotide comprising nucleotides 264-323 or 1272-1331 of the present sequence can be used as a DNA probe. The new human transferases and polynucleotides can be used in the diagnosis, prevention and treatment (including gene therapy and antisense therapy) of cancer, developmental disorders, gastrointestinal disorders, genetic disorders, immunological disorders, neurological disorders, reproductive disorders, and smooth muscle disorders.</p>					
SQ	Sequence 2065 BP; 466 A; 575 C; 491 G; 533 T; 0 other:					
	Query Match	79.5%; Score 1615; DB 21; Length 2065;				
	Best Local Similarity	99.8%; Pred. No. 0;				
	Matches 2005; Conservative	0; Mismatches 3; Indels 2; Gaps	2			
QY	9 GCCAGATGCTCCACAGTCTGGGGGAAAAAGCTTCCTCATTTCTCTTCCACCACCCTC	68				
DB	1 GCCAGATGCTCCACAGTCTGGGGGAAAAAGCTTCCTCATTTCTCTTCCACCACCCTC	60				
QY	69 AAGCAGTCCCCACCCTTGAGTCTCACAGTAGTAAAGCTGTAATTCACTTCACAGTTTC	128				
DB	61 AAGCAGTCCCCACCCTTGAGTCTCACAGTAGTAAAGCTGTAATTCACTTCACAGTTTC	120				
QY	129 TEGGAGCGAGTCTTTCTCAAGCCGCTTGCAGAAGTCTTCCACTTCACACAATGCTAC	188				
DB	121 TEGGAGCGAGTCTTTCTCAAGCCGCTTGCAGAAGTCTTCCACTTCACACAATGCTAC	180				
QY	189 TGCCATAAATAATGAAGCCTCGAGTTTCTGTTTCCAGATGGGCATCTTGGGCTAT	248				
DB	181 TGCCATAAATAATGAAGCCTCGAGTTTCTGTTTCCAGATGGGCATCTTGGGCTAT	240				
QY	249 TCTTCACATGTACAGCCACAACATCAGCTCCCTGTCTATGAAAGGCACGCCGAGCGCA	308				
DB	241 TCTTCACATGTACAGCCACAACATCAGCTCCCTGTCTATGAAAGGCACGCCGAGCGCA	300				
QY	309 TGCAGTGTGGTTCGTTCTGCTTCTGGGCGCTGGGCTCTTGTGGGGGACGTTTTG	368				
DB	301 TGCAGTGTGGTTCGTTCTGCTTCTGGGCGCTGGGCTCTTGTGTGGGGGACGTTTTG	360				
QY	369 GGCAGCACCCAGATGTTTCTACTGATGAGAGCCGCGCTGGCAGCGTGTGATGACCTTCA	428				
DB	361 GGCAGCACCCAGATGTTTCTACTGATGAGAGCCGCGCTGGCAGCGTGTGATGACCTTCA	420				
QY	429 ACCAAGCAACCGCTGAGTGTGTCACATGCGCTGTGCGGATCTGATACGGCGCTTCT	488				
DB	421 ACCAAGCAACCGCTGAGTGTGTCACATGCGCTGTGCGGATCTGATACGGCGCTTCT	480				
QY	489 TTGGGACATGAGCTCTTGTATGCTACTACATGGAACCTGTGCTCCCGGAGACAGTCCAGC	548				
DB	481 TTGGGACATGAGCTCTTGTATGCTACTACATGGAACCTGTGCTCCCGGAGACAGTCCAGC	540				
QY	549 TCTTTCAGTGGAGAACACCGCGCGCTGTGTTGTGACACTGCTGTGACATCATCCAC	608				
DB	541 TCTTTCAGTGGAGAACACCGCGCGCTGTGTTGTGACACTGCTGTGACATCATCCAC	600				
QY	609 AAGATGAATATATCCCOCGGGCTCACTGAGGCTCCTGTGCAAGTCAAAGCCCTTGG -AG	667				
DB	601 AAGATG -AATCATCCOCGGGCTCACTGAGGCTCCTGTGCAAGTCAAAGCCCTTGAAG	659				
QY	668 GTGTGGAGAAAGCGCTCGGCTCTACACACACAGTGTCTCAAGGAGTGGGCTTCTC	727				
DB	660 GTGTGGAGAAAGCGCTCGGCTCTACACACACAGTGTCTCAAGGAGTGGGCTTCTC	719				
QY	728 AACCTGAGTCCCTTACCCGCTGCTGAAGAAGCCCTCCCTCAACTGCATATGTGCAC	787				
DB	720 AACCTGAGTCCCTTACCCGCTGCTGAAGAAGCCCTCCCTCAACTGCATATGTGCAC	779				

QY	788	CTGTTCGGGAAACCCCGGCGCCGTTCCTCCCTCCCGAAGACACAAAGGGAATCTCATG	847
Db	780	CTGTTCGGGAAACCCCGGCGCGGTTCCTCCCTCCCGAAGACACAAAGGGAATCTCATG	839
QY	848	ATTGACACTGCATTTGTGATGGGACAGCATGACACAGAACTCAAGAAAGGAGACCAACCC	907
Db	840	ATTGACACTGCATTTGTGATGGGAGAGCATGACACAAACTCAAGAAAGGAGACCAACCC	899
QY	908	TACTATGTGATGCAGGTCATCTGCGAAAGGACAGCTGGAATCTACAAAGCATTCAGTC	967
Db	900	TACTATGTGATGCAGGTCATCTGCGAAAGGACAGCTGGAATCTACAAAGCATTCAGTC	959
QY	968	TTTGCCCAAGGCCCTGCAGGAAGCGTACCTCTTCGTGCGGTATGAGACCTGGCTGAGCC	1027
Db	960	TTTGCCCAAGGCCCTGCAGGAAGCGTACCTCTCTGTGCGGTATGAGACCTGGCTGAGCC	1019
QY	1028	CGTGGGGCCGAGATTCTCCGAATGTATGATTCGTGTGGGATTTGGAAATCTGTGGCCATCT	1087
Db	1020	CGTGGGGCCGAGACTTCCCAATGTATGATTTGTGTGGGATTTGGAAATCTGTGGCCATCT	1079
QY	1088	CAGACTGGGTGCTAATACATCACCCGAGGCAAGGGGATGGGTGACACGCTTCCACACA	1147
Db	1080	CAGACTGGGTGCTAATACATCACCCGAGGCAAGGGGATGGGTGACACGCTTCCACACA	1139
QY	1148	AATGCGAGGGATGGCCCTATATGTCTCCAGGGCTGTGGCGCTGGTCTTGCCCTATGAAG	1207
Db	1140	AATGCGAGGGATGGCCCTATATGTCTCCAGGGCTGTGGCGCTGGTCTTGCCCTATGAAG	1199
QY	1208	GTTCCTGCACCTGAGAAAGCCTGTGGCGATGCCATGAATTTCTGGGTCACCCGCAATC	1267
Db	1200	GTTCCTGCACCTGAGAAAGCCTGTGGCGATGCCATGAATTTCTGGGTCACCCGCAATC	1259
QY	1268	AGATCTCAACAAAGACGAGAAGAACTGTGTGGGATCTCTCTCACTCACTGAGACTGTCCCT	1327
Db	1260	AGATCTCAACAAAGACGAGAAGAACTGTGTGGGATCTCTCTCACTCACTGAGACTGTCCCT	1319
QY	1328	GAGCAAAATCCACTAAGAGGGTTAGAAAGCCTTGTGCTGCACCTGGTGTGACGCTCAGTCA	1387
Db	1320	GAGCAAAATCCACTAAGAGGGTTAGAAAGCCTTGTGCTGCACCTGGTGTGACGCTCAGTCA	1379
QY	1388	CTTCTCTCGAATGGTTCTGTGAGCCCTGCTCAACATCTGTGAGCCTTAACATACATGTCGTGG	1447
Db	1380	CTTCTCTCGAATGGTTCTGTGAGCCCTTGCCTTCATCTGTAGCCTTAACTACATAGTCTGGG	1439
QY	1448	GTAATCACTGAGTGTGATGTGTGCTCCACAGCTGCTCAAGCGAAGGACTTTGTGTCA	1507
Db	1440	GTAATCACTGAGTGTGATGTGTGCTCCACAGCTGCTCAAGCGAAGGACTTTGTGTCA	1499
QY	1508	TGCTTGTGTAGAAAACACACAGTGGGGAACCTTATGTGAGCGACATCTCCACACAGTAA	1567
Db	1500	TGCTTGTGTAGAAAACACACAGTGGGGAACCTTATGTGAGCGACATCTCCACACAGTAA	1559
QY	1568	ACAGGGATTGCTCTCTCTTTCTTTCTTGTATCTTCTCTGTCTGGGCGAGCTTCAGAGACTTT	1627
Db	1560	ACAGGGATTGCTCTCTCTCTTTCTTTCTTGTATCTTCTCTGTCTGGGCGAGACTTCAGAGACTTT	1619
QY	1628	GTGGCCTGGAGGCCATTATAGCACGACACAGATATAGTGAATTAATCCATTAACCTGCC	1687
Db	1620	GTGGCCTGGAGGCCATTATAGCACGACACAGATATAGTGAATTAATCCATTAACCTGCC	1679
QY	1688	TGTCACACATTTGGCCCAATGGGGATGTGATCTTTTCAACAAAGAGCTCACAGACTTTTCC	1747
Db	1680	TGTCACACATTTGGCCCAATGGGGATGTGATCTTTTCAACAAAGAGCTCACAGACTTTTCC	1739
QY	1748	ACAGAGATGCAAAATTCGAGACCCTTGTAGTCTCCAGGTGATTCAAGAAAGAACTGGGAA	1807
Db	1740	ACAGAGATGCAAAATTCGAGACCCTTGTAGTCTCCAGGTGATTCAAGAAAGAACTGGGAA	1799
QY	1808	CAAGGTTGATGCTCTACTATATGACTTACACATCACAGCTATCGGTAATCAGAAATATGA	1867
Db	1800	CAAGGTTGATGCTCTACTATATGACTTACACATCACAGCTATCGGTAATCAGAAATATGA	1859

OY	1868	AAAAAATCTGTGCACAAAGAGAAGCAACTTGTAGTTCACAGAGTGCCCTGGGCTCATTT	1927
XX			
CC	1869	AAAATAATCTGTGCACAAAGAGAAGCAACTTGTAGTTCACAGAGTGCCCTGGGCTCATTT	1927
DB	1860	AACAATAATCTGTGCACAAAGAGAAGCAACTTGTAGTTCACAGAGTGCCCTGGGCTCATTT	1919
XX			
OY	1928	GAAATACACTTCCCCTCGATTTCCTCCATACATAGAGACCTTTGACCTGTGAAGCTGC	1987
DB	1920	GAAATACACTTCCCCTCGATTTCCTCCATACATAGAGACCTTTGACCTGTGAAGCTGC	1979
XX			
OY	1988	CATCGTTTAATACATAAATTCGCCAATAAG	2017
DB	1980	CATCGTTTAATACATAAATTCGCCAATAAG	2009
XX			
DE	HUMAN L-selectin sulfotransferase-2 (LST-2) cDNA.		
XX			
AC	AA516947;		
DT	12-MAR-2002 (first entry)		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	111..1253	
FT		/tag= a	
FT		/product= "Human LST-2"	
XX			
FN	WO200185177-A1.		
PD	15-NOV-2001.		
XX			
PF	10-MAY-2001; 2001MO-US15452.		
XX			
PR	11-MAY-2000; 2000US-0569320.		
XX			
PA	(BURN-) BURNHAM INST.		
XX			
PI	Fukuda M, Yeh J, Hiraoaka N;		
DR	WPI; 2002-075226/10.		
XX			
PT	P-PSDB; AA011274.		
FT	New enzyme, useful for modifying acceptor molecule, comprises an isolated L-selectin sulfotransferase-2 that directs expression of L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or intestinal GlcNAc 6-sulfotransferase		
XX			
PS	ClaIm 19; Fig 4; 98pp; English.		
XX			
CC	The present invention provides a method of modifying an acceptor molecule by contacting the acceptor with an isolated betaal,3-N-acetylglucosaminyltransferase (betaal,3gnt) or an active fragment, where betaal,3gnt directs expression of a MECA-79 antigen. The invention also provides a method of treating or preventing an L-selectin-mediated condition by reducing the expression or activity of a betaal,3gnt that directs expression of a MECA-79 antigen. This can be done by administering to the subject an oligosaccharide L-selectin antagonist that inhibits binding of L-selectin to a MECA-79 antigen, for example by administering antibody material that specifically binds betaal,3gnt, and/or a betaal,3gnt antisense nucleic acid molecule. L-selectin sulfotransferase-2 (LST-2) also directs MECA-79 antigen expression. Alternatively, the expression or activity of LST-2 or its expression		

CC fragment can be reduced in combination with reducing the expression or
 CC activity of beta1,3gal. The method is useful for treating L-selectin
 CC mediated conditions such as Crohn's disease and ulcerative colitis,
 CC inflammatory disorders of the skin such as allergic contact dermatitis,
 CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type
 CC hypersensitivity reactions, diabetes and hyperplastic thymus. This
 CC sequence represents cDNA encoding human LSTR-2.

XX Sequence 1333 BP; 286 A; 393 C; 336 G; 318 T; 0 other;

Query Match 59.6%; Score 1211; DB 24; Length 1333;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1261; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

161 AAGGTCTTCACCTTCAGACACATGCTACTGCTAAAAAATGAGTCTGCTGTTCTG 220
 72 AAGGTCTTCACCTTCAGACACATGCTACTGCTAAAAAATGAGTCTGCTGTTCTG 131
 221 GTTTCACATGAGCCTCTGCTGCTATCTTCTTCCATGATGACAGCCACATCAGCTCC 280
 132 GTTTCACATGAGCCTCTGCTGCTATCTTCTTCCATGATGACAGCCACATCAGCTCC 191
 281 CTGTCTATGAGGACAGCCGAGCCGATGACAGTCTGCTGCTGCTGCTGCTGCTGCT 340
 192 CTGTCTATGAGGACAGCCGAGCCGATGACAGTCTGCTGCTGCTGCTGCTGCTGCT 251
 341 GGTCTCTCTTTGGGGGAGCTTTTGGGAGCAGCCAGATGTTTCTACCTGATGAG 400
 252 GGTCTCTCTTTGGGGGAGCTTTTGGGAGCAGCCAGATGTTTCTACCTGATGAG 311
 401 CCGGCTTCGACGCTGTGATGACCTTCAAGCAGACAGCCGCTGATGCTGACATGCT 460
 312 CCGGCTTCGACGCTGTGATGACCTTCAAGCAGACAGCCGCTGATGCTGACATGCT 371
 461 GTGGGGATCTGATAGCGGCGCTTCTTGTGACATGAGCCTTTGATGCTGCTACATG 520
 372 GTGGGGATCTGATAGCGGCGCTTCTTGTGACATGAGCCTTTGATGCTGCTACATG 431
 521 GAACCTGTGCTTCGAGAGAGTCCAGCTCTTCAAGTGGAGAGACCGGCGCTGTGT 580
 432 GAACCTGTGCTTCGAGAGAGTCCAGCTCTTCAAGTGGAGAGACCGGCGCTGTGT 491
 581 TGTGACCTGCTGTGATGACATCCTCCACAAAGATGAATCATCCCGGGGCTCACTGAG 640
 492 TGTGACCTGCTGTGATGACATCCTCCACAAAGATGAATCATCCCGGGGCTCACTGAG 551
 641 CTCTGTGACATGACAGCCCTTGTGAGGTGTGAGAGAGCTGCGCTCTCAACGCCAC 700
 552 CTCTGTGACATGACAGCCCTTGTGAGGTGTGAGAGAGCTGCGCTCTCAACGCCAC 611
 701 GTGTGCTCAAGAGGTGCGCTTCTTCAACCTGCACTGCTTCAACGCCAGTGAAGAC 760
 612 GTGTGCTCAAGAGGTGCGCTTCTTCAACCTGCACTGCTTCAACGCCAGTGAAGAC 671
 761 CCCCTCCCTCAACCTGATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
 672 CCCCTCCCTCAACCTGATGATGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731
 821 CGAGACGACAAAGGAGATCTCATGATGACAGTGTGATGAGGAGCAGATGAG 880
 732 CGAGACGACAAAGGAGATCTCATGATGACAGTGTGATGAGGAGCAGATGAG 791
 881 CAGAACTCAAGAGAGGAGCAACCTACTATGATGAGTGTGATGAGTGTGATGAGTGT 940
 792 CAGAACTCAAGAGAGGAGCAACCTACTATGATGAGTGTGATGAGTGTGATGAGTGT 851
 941 CTGAGATTCACAGACATTCATCTTCCCAAGGCGCTTCAGAGACCTACTGCTT 1000
 852 CTGAGATTCACAGACATTCATCTTCCCAAGGCGCTTCAGAGACCTACTGCTT 911
 1001 GTGGCTATGAGGAGCTGCTGAGCCCTGCTGAGCCAGCTTCCGAAATGATGATTC 1060
 912 GTGGCTATGAGGAGCTGCTGAGCCCTGCTGAGCCAGCTTCCGAAATGATGATTC 971

1061 GTGGATTTGGAATTTCTTCCCATCTTTCAGACCTGCTGCTGCTTAAATCATCACCAGGCAAG 1120
 972 GTGGATTTGGAATTTCTTCCCATCTTTCAGACCTGCTGCTGCTTAAATCATCACCAGGCAAG 1031
 1121 GGCATGGGTGACACGCTTTCACACAATGACAGAGGATGCTTAAATGCTTCCAGGCT 1180
 1032 GGCATGGGTGACACGCTTTCACACAATGACAGAGGATGCTTAAATGCTTCCAGGCT 1091
 1181 TGGCGCTGCTTTCCTTCCCTATGAAAAAGTTTCTGATCTTCAGAAAGCCTGCGGATGCC 1240
 1092 TGGCGCTGCTTTCCTTCCCTATGAAAAAGTTTCTGATCTTCAGAAAGCCTGCGGATGCC 1151
 1241 ATGATTTGCTGGGCTTACCGGACGCTGATGATGATGATGATGATGATGATGATGATGATG 1300
 1152 ATGATTTGCTGGGCTTACCGGACGCTGATGATGATGATGATGATGATGATGATGATGATG 1211
 1301 GATCTTCTGCTTACCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1360
 1212 GATCTTCTGCTTACCTGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1271
 1361 GGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1420
 1272 GGTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1331
 1421 CT 1422
 1332 CT 1333

RESULT 5
 AAK91803
 ID AAK91803 standard; cDNA; 877 BP.

AAK91803;

06-NOV-2001 (first entry)

Human cDNA 5'-end sequence, SEQ ID NO: 263.

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.

Homo sapiens.

EP1130094-A2.

05-SEP-2001.

07-JUL-2000; 2000EP-0114089.

08-JUL-1999; 99JP-0194486.

11-JAN-2000; 2000JP-0118774.

02-MAY-2000; 2000JP-0183765.

(HELI-) HELIX RES INST.

Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

WPI; 2001-52425/58.

830 Primers useful for synthesizing full length cDNA clones and their

use in genetic manipulation -

claim 2; SEQ ID NO 263; 1380bp + sequence listing; English.

The invention relates to primers for synthesizing full length cDNA

clones. 830 cDNA molecules encoding a human protein have been

isolated and nucleotide sequences of 5' and 3' ends of the cDNA

molecules have been determined. Primers for synthesizing the full length

cDNA are useful for clarifying the function of the protein encoded by

the cDNA. The full length clones were obtained by construction of full

length enriched cDNA libraries that were synthesised by the oligo-capping

CC method. the primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is the nucleotide
CC sequence of the 5'-end of a cDNA provided in the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
XX Sequence 877 BP; 183 A; 266 C; 221 G; 203 T; 4 other;

Query Match	36.5%	Score 741;	DB 22;	Length 877;
Best Local Similarity	100.0%	Pred. No. 1.1e-305;		
Matches 741; Conservative	0;	Mismatches	0;	Indels 0; Gaps 0;

QY	114	ACTTTCACAGCTTCTGTGGAGACCGAGTGCCTTCTCATAACCCGCTCTGCAAGTCTCCACT	173
Db	1	ACTTTCACAGCTTCTGTGGAGACCGAGTGCCTTCTCATAACCCGCTCTGCAAGTCTCCACT	60
QY	174	TCACCAACATGCTACTGCTCTTAAAAAATGAAAGCTCCTGCTGCTTTCTGTTCCACATGC	233
Db	61	TCACCAACATGCTACTGCTCTTAAAAAATGAAAGCTCCTGCTGCTTTCTGTTCCACATGC	120
QY	234	CCATCTTTGGGCTCTATTGCTTCCACTATGACGCCAACAAACATCACTGCTCCCTGTATATGAAG	293
Db	121	CCATCTTTGGGCTCTATTGCTTCCACTATGACGCCAACAAACATCACTGCTCCCTGTATATGAAG	180
QY	294	CACAGCCCCGAGCCGATCCAGCTGCTGGTTTGTCTTCCCTGGCCGCTGCTGCTCTCTTTTG	353
Db	181	CACAGCCCCGAGCCGATCCAGCTGCTGGTTTGTCTTCCCTGGCCGCTGCTGCTCTCTTTTG	240
QY	354	TGGGGCAGCTTTTGGGGCAGACCCCAATGTTTTTCTTCACGTATGAGAACCCCGCTGGACG	413
Db	241	TGGGGCAGCTTTTGGGGCAGACCCCAATGTTTTTCTTCACGTATGAGAACCCCGCTGGACG	300
QY	414	TGTGGATGACCTTTCAMACAGACGACCCGCTGATGCTGACATATGGCTGTGCGGAGATCTGA	473
Db	301	TGTGGATGACCTTTCAMACAGACGACCCGCTGATGCTGACATATGGCTGTGCGGAGATCTGA	360
QY	474	TAGGGCGCGTCTCTTGTGGACATATGAGCTCTTTGATGCTCTACATGGAACCTGTGTCCC	533
Db	361	TAGGGCGCGTCTCTTGTGGACATATGAGCTCTTTGATGCTCTACATGGAACCTGTGTCCC	420
QY	534	GGAGACAGTCCAGGCTCTTTCACTGGGAGAACAGCCGGCCCTGTGTCTGCACCTGCGCT	593
Db	421	GGAGACAGTCCAGGCTCTTTCACTGGGAGAACAGCCGGCCCTGTGTCTGCACCTGCGCT	480
QY	594	GTCGACATCATCCCAAGATGAAATATATCCCCGGGAGCTCACTGCAGGCTCTCTGTGCAATC	653
Db	481	GTCGACATCATCCCAAGATGAAATATATCCCCGGGAGCTCACTGCAGGCTCTCTGTGCAATC	540
QY	654	AACAGCCCTTTGAGTGTGTGGAGAAGGCGCTGCCGCTCTACAGCCACAGTGTGCTCAAG	713
Db	541	AACAGCCCTTTGAGTGTGTGGAGAAGGCGCTGCCGCTCTACAGCCACAGTGTGCTCAAG	600
QY	714	AGGTGGCCTTCTTCACTGCTGCAATCCCTCTTACCCGCTGCTGGAAGAAGACCCCTCCCTCAAC	773
Db	601	AGGTGGCCTTCTTCACTGCTGCAATCCCTCTTACCCGCTGCTGGAAGAAGACCCCTCCCTCAAC	660
QY	774	TGCATATCGTGACACTGTGTCCGGAGACCCCGGGCCGCTGTCCGTTCCCGGAAGACGACAA	833
Db	661	TGCATATCGTGACACTGTGTCCGGAGACCCCGGGCCGCTGTCCGTTCCCGGAAGACGACAA	720
QY	834	AGGAGATCTCATATGATTGACA	854
Db	721	AGGAGATCTCATATGATTGACA	741

DE	Human cDNA clone representative sequence, SEQ ID NO: 2381.
XX	
KW	Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX	
OS	Homo sapiens.
XX	
FN	EP130094-A2.
XX	
PD	05-SEP-2001.
XX	
FE	07-JUL-2000; 2000EP-0114089.
XX	
PR	08-JUL-1999; 99JP-0194486.
XX	
PR	11-JAN-2000; 2000JP-0118774.
XX	
PR	02-MAY-2000; 2000JP-0185765.
XX	
PA	(HELI-) HELIX RES INST.
XX	
PI	Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
PL	Makamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H
XX	
DR	WPI; 2001-524255/58.
XX	

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation -

CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence was used as the
CC representative sequence from a human clone which was used in
CC homology searches to identify the clone.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

Sequence 877 BP; 183 A; 266 C; 221 G; 203 T; 4 other;

```
Query Match      36.5%; Score 741; DB 22; Length 877;
Best Local Similarity 100.0%; Pred. No. 1.1e-305;
Matches 741; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

QY	114	ACTTCACAGCTTCTCGGACGAGGCGTTCTTCACACCGCTTTGCAAGSTCTTCCACT	173
Db	1	ACTTTCACAGCTTCTCGGACGAGGCGTTCTTCACACCGCTTTGCAAGSTCTTCCACT	60
QY	174	TCACGACACATGCTACTGCGCTTAAAAAAATGAAAGCTCTGCTGTTCTGTTTCCAGATGG	233
Db	61	TCACACACATGCTACTGCGCTTAAAAAAATGAAAGCTCTGCTGTTCTGTTTCCAGATGG	120
QY	234	CCATCTTGGCTCTATCTTCTCCACATGTAACACCCACAACATAGTGGCTGTCTATAGAGG	293
Db	121	CCATCTTGGCTCTATCTTCTCCACATGTAACACCCACAACATAGTGGCTGTCTATAGAGG	180
QY	294	CACAGCCCGAGCGCATGACAGCTGCTGTTCTGTCTTCTCTGGCGCTTGGCTCTTTCTTTG	353
Db	181	CACAGCCCGAGCGCATGACAGCTGCTGTTCTGTCTTCTCTGGCGCTTGGCTCTTTCTTTG	240
QY	354	TGGGGCAGCTTTTGGGCGACGACCCCAATGCTTTTCTACGTATGAGACCCGCGCGACG	413
Db	241	TGGGGCAGCTTTTGGGCGACGACCCCAATGCTTTTCTACGTATGAGACCCGCGCGACG	300
QY	414	TGTGGATGACCTTTCACACAGACACCCCGCTGTGATGCTGTGCACATGCTGTGCGGGATTTGA	473
Db	301	TGTGGATGACCTTTCACACAGACACCCCGCTGTGATGCTGTGCACATGCTGTGCGGGATTTGA	360
QY	474	TACGGGCGGTTCTTCTTGGACATAGAGCGTCTTGTGATGCTACATAGTAAGCACTGGTCCC	533

Db	361	TAGGGCCGCTCTCTGTGTGAGCAATGAGAGGTCTTTGATGGCCCTACATGGAACCTGTGCCCC	420
Oy	534	GGAAGACAGTCCACCCCTCTTTCACTAGTGGGAGAACAGCCGGGCCCTGTGTCTCACCCTGCT	593
Db	421	GGAGACAGTCCACCCCTCTTTCACTAGTGGGAGAACAGCCGGGCCCTGTGTCTCACCCTGCT	480
Oy	594	GTGACATCATCCACACAGATGAAATCATCCCCGGGGCTCACTGACAGGCTCTGTGCAGTC	653
Db	481	GTGACATCATCCACACAGATGAAATCATCCCCGGGGCTCACTGACAGGCTCTGTGCAGTC	540
Oy	654	AACAGCCCTTTGAGTGTGTGGAGAGAGGCGCTGCCTCTCTACAGCACAGTGTGCTCAAG	713
Db	541	AACAGCCCTTTGAGTGTGTGGAGAGAGGCGCTGCCTCTCTACAGCACAGTGTGCTCAAG	600
Oy	714	AGGAGGCGCTTCTTCACACCTGCAGAGTCCCTCTACCCGCTGTGTGAAGAACCCCTCCCAAC	773
Db	601	AGGAGGCGCTTCTTCACACCTGCAGAGTCCCTCTACCCGCTGTGTGAAGAACCCCTCCCAAC	660
Oy	774	TGCATATCGTGCACCTGTGTCCGGGAGCCCGGGCGGTTCGTTCCGAGAACGCACAA	833
Db	661	TGCATATCGTGCACCTGTGTCCGGGAGCCCGGGCGGTTCGTTCCGAGAACGCACAA	720
Oy	834	AGGAGATCTCATGATTGACA	854
Db	721	AGGAGATCTCATGATTGACA	741
RESULT 7			
XX	ABK54724		
XX	ID	ABK54724 standard; cDNA; 517 BP.	
AC	ABK54724		
XX	18-JUN-2002 (first entry)		
DT			
XX	Human colon cancer-associated cDNA, SEQ ID No 194.		
DE			
XX	Human; colon cancer; immunogenic; vaccine; tumour; gene; ss.		
KW			
XX	Homo sapiens.		
OS			
XX	W0200212280-A2.		
PN			
PD	14-FEB-2002.		
XX			
XX	30-JUL-2001; 2001WO-US23826.		
FP			
XX	03-AUG-2000; 2000US-22326SP.		
PR	02-OCT-2000; 2000US-237406P.		
PR	20-MAR-2001; 2001US-277485P.		
PR	03-JUL-2001; 2001US-302702P.		
XX			
PA	(CORI.) CORIXA CORP.		
XX			
PI	PyLe RA, Xu J, Secrist H;		
XX	WPI; 2002-257462/30.		
DR			
XX			
XX	Novel polynucleotide encoding colon tumour polypeptides, useful as		
PT	vaccines for treating colon cancers		
XX			
PS	Claim 1; Page 206; 425pp; English.		
XX			
CC	The invention relates to isolated polynucleotides (I) encoding colon		
CC	tumour polypeptides (II). (I) is useful for stimulating an immune		
CC	response in a patient and treating colon cancer in a patient.		
CC	Oligonucleotides derived from (I) are useful for determining the presence		
CC	of cancer in a patient. (I) and (II) are useful in pharmaceutical		
CC	compositions, e.g. vaccines, and other compositions for the diagnosis		
CC	and treatment of colon cancer. A composition comprising a first component		
CC	selected from physiologically acceptable carriers and immunostimulants,		
CC	and an antigen-presenting cell expressing (II) is useful for inhibiting		

CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). AB54531-AB55464 represent human colon cancer cDNA
CC sequences of the invention.
xx
SQ Sequence 517 BP; 121 A; 130 C; 124 G; 142 T; 0 other;

```
Query Match      24.9%; Score 505; DB 24; Length 517;
Best Local Similarity 100.0%; Pred. No. 3.8e-205;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1355 GGCCTTGGTCGCCACTGGGTGTCAGCCTCAATCACTTTCTCTGAATGCTTCTGAGCCTTGC 14
|||||
Db 13 GGCCTTGGTCGCCACTGGGTGTCAGCCTCAATCACTTTCTCTGAATGCTTCTGAGCCTTGC 72

Qy 1415 CTGATCTCTAGCCTTAAGTACATGCTGTGGGATACAGCTGAGTGTGCTCC 1474

QY 1475 ACACGCTCAAGCAGAAGACTTTTGTGTCCATGCTGTGTCTGAAAAACAGACTGGG 1534

QY 1535 AACCTATGTGAGCAGACATCCCAACAGTGAACAGGGTATTGCTCTCTCTTTCTT 1594

1505 CATCTTCTCTCTCTGGGCAAGACCTTTCAGACACCTTTGCTGGGCTTGGAGAGGCTATTAAAGCAGGAC 1552
Db 193 AACCTATATGTGAGCAGACATATCCCAACAGTGTAAACAGGGATATGCTCTCTCTCTCTCTT 252

Db 253 GATCTTCCTCTCTGGGCAGACTTCAGAGACTTTGTGSCCTGGAGGCCTATTAAAGCAGAC 312

Db 313 ACAGTACAGTGGAAATGATCCATAAACCCTCCCTGCCACATCTGCCCAATGGGGAATG 372

QY 1715 GATCTTTCACCAAGAGTCACCAAGATTTTCCACAGAGATGCAAAATTCTGAGCCCTTGG 1774

QY 1775 AGTCCAGTGGATTCAAGGAGGAGTGGGACACAGGTTGGATGCCCTACTATGAGCTT 183

QY 1835 GACCATCACGCTATCGTATCAG 1859

Db 493 GACCATCACAGCTATCGSTATACG 517

TD AAK92588 standard: CDNA: 548 BP.
AAK92588/c
RESULT 8

AAK9258;
AC
XX

XX
DE Human cDNA 3'-end sequence, SEQ ID NO: 1048.

Human; full length cDNA; cDNA synthesis; oligo-capping; ss.
XX
KW
XX
Homo sapiens
OS

AA
PN
XX
EP1130094-A2.

XX 07-JUL-2000; 2000EP-0114089.
PF

PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 08-MAY-2000; 2000JD-0182755

XX
PA (HELI-) HELIX RES INST.

XX Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y,
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX WPI; 2001-524255/58.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 3; SEQ ID NO 1048; 1380bp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 nucleotide sequences encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is the nucleotide
 CC sequence of the 3'-end of a cDNA provided in the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 CC
 XX
 SQ Sequence 548 BP; 160 A; 109 C; 129 G; 148 T; 2 other;

Query Match 17.3%; Score 351; DB 22; Length 548;
 Best Local Similarity 100.0%; Pred. No. 1.5e-139;

Matches 351; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1667 GATTGATCCATAAACCCTCCCTGTCACATCTTGCCCAATGGGGAATGATCTTACCA 1726
 DB 426 GAATGATCCATAAACCCTCCCTGTCACATCTTGCCCAATGGGGAATGATCTTACCA 367
 OY 1727 AAGACTCACCCGATTTTCCACAGATGCAATTCGAGCCCTTGAGTTCCCAATGG 1786
 DB 366 AAGACTCACCCGATTTTCCACAGATGCAATTCGAGCCCTTGAGTTCCCAATGG 307
 OY 1787 ATTCAAG 1846
 DB 306 ATTCAAG 247
 OY 1847 TTTCGCTATGCAAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 1906
 DB 246 TTTCGCTATGCAAAATGGAATGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 187
 OY 1907 CAGGCTGCTGGGCTGCAATTTGAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 1966
 DB 186 CAGGCTGCTGGGCTGCAATTTGAATGCAATTTGCAATTTGCAATTTGCAATTTGCA 127
 OY 1967 GACTTTGACCTGTGAAGCTGCTGTGTAATGTAATGTAATGTAATGTAATGTAATG 2017
 DB 126 GACTTTGACCTGTGAAGCTGCTGTGTAATGTAATGTAATGTAATGTAATGTAATG 76

RESULT 9
 ID AAC76156
 AAC76156 standard; cDNA; 2988 BP.

AC AAC76156;

DT 08-FEB-2001 (first entry)

DE Human ORF1711 polynucleotide sequence SEQ ID NO:3421.

XX Human; open reading frame; ORF; detection; cytosolic; hepatotropic;
 KW vulnery; antipapillary; antipapillary; antipapillary; antipapillary;
 KW anticonvulsant; osteoporotic; antipapillary; immunosuppressant; candidant;
 KW immunosuppressant; thrombolytic; coagulant; vasodilator; antidiabetic;
 KW hypotensive; dermatologic; immunosuppressive; antineoplastic;
 KW antiviral; antibacterial; antifungal; antineoplastic; antithyroid;
 KW antineoplastic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;

KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200058473-A2.
 XX
 PD 05-OCT-2000.
 XX
 PP 31-MAR-2000; 2000WO-US08621.
 PP
 PR 31-MAR-1999; 99US-0127607.
 PR 02-APR-1999; 99US-0127636.
 PR 05-APR-1999; 99US-0127728.
 PR 30-MAR-2000; 2000US-0540763.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Shinketsu RA, Leach M;
 XX
 DR WPI; 2000-602362/57.
 DR P-PSDB; AAB41947.
 XX
 XX Novel nucleic acids and peptides derived from open reading frame X,
 PT useful for treating e.g. cancers, proliferative disorders,
 PT neurodegenerative disorders and cardiovascular disease -
 XX
 Claim 5; Page 2597-2599; 5507pp; English.

AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 which represent the human ORF open reading frames 1 to 3161. The ORF
 sequences have activities such as: cytosolic; hepatotropic; vulnery;
 CC antipapillary; antipapillary; antipapillary; antipapillary;
 CC osteoporotic; anticonvulsant; antipapillary; immunosuppressant;
 CC immunosuppressant; candidant; thrombolytic; coagulant; vasodilator;
 CC antidiabetic; hypotensive; dermatologic; immunosuppressive;
 CC antineoplastic; antipapillary; antipapillary; antipapillary;
 CC antithyroid; antineoplastic. The sequences can be used for determining
 CC the presence of or predisposition to, or preventing or treating
 CC pathological conditions associated with an ORF-associated disorder. The
 CC nucleic acids can be used to express ORF proteins in gene therapy
 CC vectors. The proteins and nucleic acids may be used to treat cancers,
 CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
 CC graft vs host disease, cardiovascular disease, diabetes mellitus,
 CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
 CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
 CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
 CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
 CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance
 CC coagulation; to inhibit thrombosis; and as a contraceptive.

SO Sequence 2988 BP; 613 A; 916 C; 842 G; 617 T; 0 other;

Query Match 15.1%; Score 306; DB 21; Length 2988;
 Best Local Similarity 100.0%; Pred. No. 2e-120;

Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 506 TTGATGCTATGATGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 565
 DB 1480 TTGATGCTATGATGAGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1539
 OY 566 AGCCGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 625
 DB 1540 AGCCGGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1599
 OY 626 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 685
 DB 1600 CGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1659

QY 686 CGCTCTACAGCCAGCGTGGTCTCAAGAGGTGGCTTTCTTCAACCGTGCAGTCCCTTAC 745
DB 1660 CGCTCTACAGCCAGCGTGGTCTCAAGAGGTGGCTTTCTTCAACCGTGCAGTCCCTTAC 1719
QY 746 CCGCTGTGAAGACCCCTCCCTCAACCTCATATGCTGCACCTGGTCCGGACCCCGG 805
DB 1720 CCGCTGTGAAGACCCCTCCCTCAACCTCATATGCTGCACCTGGTCCGGACCCCGG 1779
QY 806 GCGGTG 811
DB 1780 GCGGTG 1785

RESULT 10
ABK54794
ID ABK54794 standard; cDNA; 389 BP.
AC ABK54794;
XX
DT 18-JUN-2002 (first entry)
XX
DE Human colon cancer-associated cDNA, SEQ ID NO 264.
XX
KM Human: colon cancer; immunogenic; vaccine; tumour; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200212280-A2.
XX
PD 14-FEB-2002.
XX
PF 30-JUL-2001; 2001WO-0523826.
XX
PR 03-AUG-2000; 2000US-223265P.
PR 02-OCT-2000; 2000US-237406P.
PR 20-MAR-2001; 2001US-277455P.
PR 03-JUL-2001; 2001US-302702P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Pyle RA, Xu J, Secrist H;
XX
DR WPI; 2002-257462/30.
XX
PT Novel polynucleotide encoding colon tumour polypeptides, useful as
PT vaccines for treating colon cancers
XX
PS Claim 1; Page 225; 425pp; English.
XX
CC The invention relates to isolated polynucleotides (I) encoding colon
CC tumour polypeptides (II). (I) is useful for stimulating an immune
CC response in a patient and treating colon cancer in a patient.
CC Oligonucleotides derived from (I) are useful for determining the presence
CC of cancer in a patient. (I) and (II) are useful in pharmaceutical
CC compositions, e.g. vaccines, and other compositions for the diagnosis
CC and treatment of colon cancer. A composition comprising a first component
CC selected from physiologically acceptable carriers and immunostimulants,
CC and an antigen-presenting cell expressing (II) is useful for inhibiting
CC development of cancer in a patient. (I) is useful in the design and
CC preparation of ribozyme molecules for inhibiting expression of tumour
CC polypeptides and (I). ABK54531-ABK5464 represent human colon cancer cDNA
CC sequences of the invention.
SQ Sequence 389 BP; 121 A; 83 C; 76 G; 109 T; 0 other;

Query Match 14.18; Score 286; DB 24; Length 389;
Best Local Similarity 100.0%; Pred. No. 7.5e-112;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1732 CTCACGACATTTCCACAGATGCAAAATTCGAGCCTTGAGTTCCAGTGATTC 1791
DB 13 CTCACGACATTTCCACAGATGCAAAATTCGAGCCTTGAGTTCCAGTGATTC 72

QY 1792 AGGAAGAGTGGGAACAAGTTGGATGCTACTTATGAGCTTGACATCAGCTATCG 1851
DB 73 AGGAAGAGTGGGAACAAGTTGGATGCTACTTATGAGCTTGACATCAGCTATCG 132
QY 1852 GTAATCAGAAATATGAAACAAATCTCTGCACAAAAGCAAGCTTTAAGTTCACAGG 1911
DB 133 GTAATCAGAAATATGAAACAAATCTCTGCACAAAAGCAAGCTTTAAGTTCACAGG 192
QY 1912 TGCTGGGCTGATTTGAATACATCTCCCTCTGATTTTCCATCATAGAACTT 1971
DB 193 TGCTGGGCTGATTTGAATACATCTCCCTCTGATTTTCCATCATAGAACTT 252
QY 1972 TGACCTGTGAAGCTGCATCTGTATATCTAAATTCCTCAATAG 2017
DB 253 TGACCTGTGAAGCTGCATCTGTATATCTAAATTCCTCAATAG 298

RESULT 11
ABN38074
ID ABN38074 standard; DNA; 60 BP.
AC ABN38074;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human spliced transcript detection oligonucleotide SEQ ID NO:10822.
XX
KM Human; mouse; rat; splice transcript; detection; RNA transcript;
KM splice variant; transcriptome; oligonucleotide library; ss.
XX
OS Homo sapiens.
XX
PN WO200210449-A2.
XX
PD 07-FEB-2002.
XX
PF 20-JUL-2001; 2001WO-1B01903.
XX
PR 28-JUL-2000; 2000US-221607P.
PR 02-MAY-2001; 2001US-287724P.
XX
PA (COMP-) COMPUGEN INC.
XX
PI Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
XX
DR WPI; 2002-257383/30.
XX
PT New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
XX
PS Example 1; SEQ ID 10822; 47pp; English.
XX
CC The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridizing selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterizing the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mint
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN9589 represent

CC oligonucleotide sequences from rats, humans and mice, which are used in
 CC the exemplification of the present invention.
 CC N.B. The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 60 BP; 13 A; 20 C; 9 G; 18 T; 0 other;
 SQ

Query Match 3.0%; Score 60; DB 24; Length 60;
 Best Local Similarity 100.0%; Pred. No. 1.5e-15;
 Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1934 CACTGCCCTGTCATTTCCATCAGATGAAAGACTTGGACCTGTGAAGCTGCCATCTG 1993
 DB 1 CACTGCCCTGTCATTTCCATCAGATGAAAGACTTGGACCTGTGAAGCTGCCATCTG 60

RESULT 12
 AAD24670
 ID AAD24670 standard; CDNA; 1647 BP.
 XX
 XX AAD24670;
 AC
 AC 12-MAR-2002 (first entry)
 DT
 DT Human drug metabolizing enzyme (DME)-5 CDNA.
 DE
 XX Human; drug metabolizing enzyme; gene therapy; autoimmune disorder;
 XX inflammatory disorder; acquired immune deficiency syndrome; infection;
 XX AIDS; adult respiratory distress syndrome; cell proliferative disorder;
 XX allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;
 XX cancer; endocrine disorder; hypothalamus disorder; pituitary disorder;
 XX gastrointestinal disorder; metabolic disorder; developmental disorder;
 XX liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 XX goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;
 XX DME-5; ss.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT 1..1188
 FT CDS /tag= a
 FT /product= "Human DME-5 protein"
 FT sig_peptide 1..96
 FT /tag= b
 FT mat_peptide 97..1185
 FT /tag= c
 FT /product= "Human mature DME-5 protein #1"
 FT sig_peptide 1..105
 FT /tag= d
 FT mat_peptide 106..1185
 FT /tag= e
 FT /product= "Human mature DME-5 protein #2"
 FT
 XX
 XX WO200179468-A2.
 PN
 PN 25-OCT-2001.
 PD
 PD 12-APR-2001; 2001WO-US11869.
 PF
 PF 13-APR-2000; 2000US-197590P.
 PR 19-APR-2000; 2000US-198403P.
 PR 28-APR-2000; 2000US-200185P.
 PR 05-MAY-2000; 2000US-202234P.
 PR 11-MAY-2000; 2000US-203509P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 PA Policky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM,
 PI Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS,
 PI Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 PI Au-Young J;
 XX

DR WPI: 2002-066363/09.
 DR P-PSDB: ABE15438.
 XX
 XX
 XX Novel isolated human drug metabolizing enzymes referred as DME 1-10,
 PT useful for diagnosing, treating, or preventing disorders associated
 PT with aberrant expression of DME such as allergy, anemia, asthma,
 PT infertility
 XX
 XX Claim 5; Page 139; 143pp; English.
 XX
 XX The invention relates to human drug metabolizing enzymes referred as
 CC DME and nucleic acid molecules encoding such enzymes. Polynucleotides
 CC of the invention are useful for assessing toxicity of test compounds
 CC and in gene therapy. Sequences of the invention are useful in the
 CC diagnosis, prevention and treatment of autoimmune/inflammatory
 CC disorders such as acquired immune deficiency syndrome (AIDS), adult
 CC respiratory distress syndrome, allergies, anaemia, arteriosclerosis,
 CC asthma, autoimmune haemolytic anaemia, contact dermatitis, Crohn's
 CC disease, glomerulonephritis, Goodpasture's syndrome, gout, Grave's
 CC disease, Hashimoto's thyroiditis, irritable bowel syndrome, multiple
 CC sclerosis, osteoarthritis, osteoporosis, psoriasis, systemic lupus
 CC erythematosus, rheumatoid arthritis, scleroderma, ulcerative colitis,
 CC uveitis, viral, bacterial, fungal, parasitic, protozoal, helminthic
 CC infections; cell proliferative disorders such as actinic keratosis,
 CC arteriosclerosis, atherosclerosis, Duchenne and Becker dystrophy,
 CC hepatitis, Cushing's syndrome, cancers, myelodysplastic syndrome,
 CC epilepsy; endocrine disorders such as disorders of the hypothalamus
 CC and pituitary resulting from lesions such as primary brain tumours,
 CC adenomas, infarction associated with pregnancy, aneurysms, vascular
 CC malformations; eye disorders such as conjunctivitis, iritis, retinitis,
 CC glaucoma, pigmentosa; metabolic disorders such as Addison's disease,
 CC cystic fibrosis, diabetes, goitre, glycogen storage diseases,
 CC hypercholesterolaemia, hyperthyroidism, hypoglycaemia, lipid myopathies,
 CC Meckles syndrome, mannosidosis, obesity; gastrointestinal disorders such
 CC as dysphagia, gastric carcinoma, anorexia, nausea, gastroenteritis,
 CC hyperbilirubinaemia, emesis, cirrhosis, diarrhoea, jaundice, Reye's
 CC syndrome, peliosis hepatitis, hepatic vein thrombosis and developmental
 CC disorders. The present sequence is human DME-5 protein CDNA.
 CC
 XX
 XX Sequence 1647 BP; 287 A; 560 C; 501 G; 299 T; 0 other;
 SQ

Query Match 2.9%; Score 59; DB 24; Length 1647;
 Best Local Similarity 100.0%; Pred. No. 3.4e-15;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGGCGTGGCGTCTCTACAGCCAGCGTGTGTCAGAGAGTGGCTTCTCAACCTGCAG 736
 DB 488 AGGCGTGGCGTCTCTACAGCCAGCGTGTGTCAGAGAGTGGCTTCTCAACCTGCAG 546

RESULT 13
 AAD027700
 ID AAD027700 standard; CDNA; 1694 BP.
 XX
 XX AAD027700;
 AC
 AC 02-MAY-2001 (first entry)
 DT
 DT Human glycosyl sulfotransferase-4beta (GST-4beta) cDNA.
 DE
 DE Human glycosyl sulfotransferase-4beta (GST-4beta) cDNA.
 XX
 XX Human; glycosyl sulfotransferase-4beta; GST-4beta; immunosuppressive;
 XX therapy; selectin binding inhibitor; gene therapy; inflammation;
 XX systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 XX polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 XX glomerulonephritis; masthenia gravis; Sjogren's syndrome; adrenailitis;
 XX Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
 XX demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 XX myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 XX asthma; hypersensitivity; rheumatic fever; tissue rejection;
 XX chromosome 16q23.1; ss.
 XX
 OS Homo sapiens.
 XX


```

FH Key Location/Qualifiers
FT 5'UTR 9..188
FT /*tag= a
FT CDS 189..1376
FT /*tag= b
FT /product= "Human glycosyl sulfotransferase-4beta
FT (GST-4beta)"
FT /note= "CDS is referred as SEQ ID NO:21 in brief
FT description of the figures (page no: 4)"
FT 3'UTR 1377..1694
FT /*tag= c
XX
XX MO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX (RBGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX WPI: 2001-138471/14.
XX P-PSDB; AAY72640.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Claim 6; Fig 4A; 128bp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4beta (GST-4
XX beta) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation.
XX
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX Sequence 1694 BP; 259 A; 586 C; 514 G; 335 T; 0 other;
XX
XX Query Match 2.9%; Score 59; DB 22; Length 1694;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-15;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 678 AGGCGTGGCGCTCTACAGCCAGCGTGGTCTCAAGAGGTCGCTTCTCAACCTGCAG 736
XX |||||||
XX Db 676 AGGCGTGGCGCTCTACAGCCAGCGTGGTCTCAAGAGGTCGCTTCTCAACCTGCAG 734
XX
XX RESULT 14
XX AAD02699
XX ID AAD02699 standard; cDNA; 2044 BP.
XX
XX AAD02699;
XX
XX 02-MAY-2001 (first entry)
XX
XX Human glycosyl sulfotransferase-4alpha (GST-4alpha) cDNA #2.
XX
XX Human; glycosyl sulfotransferase-4alpha; GST-4alpha; immunosuppressive;
XX
```

```

XX therapy; selectin binding inhibitor; gene therapy; inflammation;
XX systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
XX polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
XX glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
XX Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
XX demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
XX myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
XX asthma; hypersensitivity; rheumatic fever; tissue rejection;
XX chromosome 16q23.1; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 218..1390
XX /*tag= a
XX /product= "Human glycosyl sulfotransferase-4alpha
XX (GST-4alpha)"
XX /note= "CDS is specifically claimed as SEQ ID NO: 4
XX in claim 6 (page no: 41) of the specification"
XX
XX MO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX (RBGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX WPI: 2001-138471/14.
XX P-PSDB; AAY72639.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Claim 6; Fig 1; 128bp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
XX alpha) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation.
XX
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX Note: The present sequence is also shown in sequence listing (page
XX no: 56) but lacks four nucleotides at its 3' end.
XX
XX Sequence 2044 BP; 422 A; 645 C; 570 G; 407 T; 0 other;
XX
XX Query Match 2.9%; Score 59; DB 22; Length 2044;
XX Best Local Similarity 100.0%; Pred. No. 3.3e-15;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 678 AGGCGTGGCGCTCTACAGCCAGCGTGGTCTCAAGAGGTCGCTTCTCAACCTGCAG 736
XX |||||||
XX Db 708 AGGCGTGGCGCTCTACAGCCAGCGTGGTCTCAAGAGGTCGCTTCTCAACCTGCAG 766
XX
XX RESULT 15
XX
```


AAD02698
ID AAD02698 standard; cDNA: 2170 BP.
XX
AC AAD02698;
XX
DT 02-MAY-2001 (first entry)
XX
DE Human glycosyl sulfotransferase-4alpha (GST-4alpha) cDNA #1.
XX
KW Human: glycosyl sulfotransferase-4alpha; GST-4alpha; immunosuppressive;
KW therapy: selectin binding inhibitor; gene therapy; inflammation;
KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
KW chromosome 16q23.1; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX 5'UTR 9...343
XX FT /*tag= a
XX FT 344...1516
XX FT /*tag= b
XX FT /product= "Human glycosyl sulfotransferase-4alpha
XX FT (GST-4alpha)"
XX FT /note= "CDS is specifically claimed as SEQ ID NO: 4
XX FT in claim 6 (page no: 41) of the specification"
XX FT 1517..2134
XX FT 3'UTR /*tag= c
XX
XX WO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX PR 13-JUL-2000; 2000US-0593828.
XX
XX (REGC) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX WPI: 2001-138471/14.
XX DR P-PSDB; AAY72639.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX PT diagnostic and therapeutic agent screening applications
XX PT
XX
XX Claim 6; Page 62; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
XX alpha) cDNA. GST-4 gene is found on chromosome 16q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation,
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX

SO Sequence 2170 BP; 449 A; 682 C; 612 G; 427 T; 0 other;
Query Match 2.9%; Score 59; DB 22; Length 2170;
Best Local Similarity 100.0%; Pred. No. 3.3e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 678 AGGCGTCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 736
DB 834 AGGCGTCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 892
RESULT 16
ABN89506
ID ABN89506 standard; cDNA: 2544 BP.
XX
AC ABN89506;
XX
DT 05-SEP-2002 (first entry)
XX
XX Human corneal N-acetylglucosamine-6-sulfotransferase cDNA SEQ ID NO:1.
XX
KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme; GLCNAC6ST.
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW opthalmological; chromosome 16q22; gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX FT 693..1880
XX FT /*tag= a
XX FT /product= "N-acetylglucosamine-6-sulfotransferase"
XX FT
XX US2002061562-A1.
XX
XX 23-MAY-2002.
XX
XX 09-AUG-2001; 2001US-0927602.
XX
XX 11-AUG-2000; 2000US-325773P.
XX
XX (FUKU/) FUKUDA M N.
XX PA (AKAM/) AKAMA T O.
XX
XX FUKUDA MN, Akama TO;
XX WPI: 2002-507643/54.
XX DR P-PSDB; ABB81554.
XX
XX New nucleic acid encoding corneal
XX PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
XX PT monitoring and diagnosis of macular corneal dystrophy
XX PT
XX Claim 4; Fig 1A-D; 69pp; English.
XX
XX The present sequence encodes human corneal
XX CC N-acetylglucosamine-6-sulfotransferase (1), which is able to catalyze
XX CC sulfation of keratan sulfate (KS). Also described is a method for
XX CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
XX CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
XX CC and has ophthalmological activity. (I) can be used to treat or prevent
XX CC macular corneal dystrophy types I or II. (I) makes possible treatment
XX CC of MCD without requiring keratoplasty or keratectomy.
XX
XX Sequence 2544 BP; 460 A; 800 C; 733 G; 551 T; 0 other;
QY 678 AGGCGTCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 736
DB 1180 AGGCGTCGCGCTCTACAGCCAGCGTGTGCTCAAGAGGTCGCTTTCACCTGCAG 1338
Query Match 2.9%; Score 59; DB 24; Length 2544;
Best Local Similarity 100.0%; Pred. No. 3.3e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

FT 3'UTR 99662..99968
FT /*tag= t
XX
XX MO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 99US-0144694.
XX 13-JUL-2000; 2000US-0593828.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX
XX WPI; 2001-138471/14.
XX P-PSDB; AAY72639, AAY72640.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications
XX
XX Example 1; Page 62-104; 128pp; English.
XX
XX The present sequence is human glycosyl sulfotransferase-4 (GST-4) genomic
XX DNA encoding GST-4alpha and GST-4beta. GST-4 gene is found on
XX chromosome 10q23.1.
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand, GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation,
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, periculous
XX anaemia, demyelinating diseases, cirrhosis, uicerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX Sequence 160552 BP; 40281 A; 37573 C; 38015 G; 44564 T; 119 other;
XX
XX Query Match 2.9%; Score 59; DB 22; Length 160552;
XX Best Local Similarity 100.0%; Pred. No. 2.5e+15;
XX Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 678 AGGCGTGGCGCTCTACAGCCAGCGTGTGCTCAAGGAGCGTGGCTTCTCAACCTGCAG 736
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 48446 AGGCGTGGCGCTCTACAGCCAGCGTGTGCTCAAGGAGCGTGGCTTCTCAACCTGCAG 48504
XX
XX RESULT 19
XX AA220798
XX ID AA220798 standard; DNA; 37 BP.
XX
XX AA220798;
XX
XX 08-DEC-1999 (first entry)
XX
XX PCR primer for glycosyl sulfotransferase-3 coding sequence.
XX
XX DE Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
XX KW selectin binding interaction; inflammation; lymphocyte homing; human;
XX KW secondary lymph organ; ss.
XX
XX OS Synthetic.
XX
XX PN MO9949018-A1.
XX
XX PD 30-SEP-1999.

```

```

XX
XX 26-FEB-1999; 99WO-US04316.
XX
XX 20-MAR-1998; 98US-0045284.
XX 12-NOV-1998; 98US-0190911.
XX
XX (REGC ) UNIV CALIFORNIA.
XX (SYNT ) SYNTAX USA INC.
XX
XX Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
XX
XX WPI; 1999-580442/49.
XX
XX Human and murine glycosyl sulfotransferase 3 and related
XX polynucleotides
XX
XX Example 4; Page 30; 59pp; English.
XX
XX This sequence encodes the human glycosyl sulfotransferase-3 (GST-3) of
XX the invention. The nucleic acid sequences, probes and primers derived
XX from these, proteins and antibodies are useful in detecting homologues.
XX The sequences, antibodies and methods are useful in the diagnosis and
XX treatment of diseases associated with selectin binding interactions,
XX including conditions associated with or resulting from the homing of
XX leukocytes to sites of inflammation and the normal homing of lymphocytes
XX to secondary lymph organs.
XX
XX SQ Sequence 37 BP; 14 A; 9 C; 8 G; 6 T; 0 other;
XX
XX Query Match 1.8%; Score 37; DB 20; Length 37;
XX Best Local Similarity 100.0%; Pred. No. 9.9e-06;
XX Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 884 AACTCAAGAAGAGGACCAACCTACTATGTATGTC 920
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 1 AACTCAAGAAGAGGACCAACCTACTATGTATGTC 37
XX
XX RESULT 20
XX ABR70821
XX ID ABR70821 standard; DNA; 30 BP.
XX
XX ABR70821;
XX
XX 15-JUL-2002 (first entry)
XX
XX DE Probe for human gene CHST4.
XX
XX Human; sulphuric acid conjugation; ss; probe; CHST1; CHST3;
XX KW CHST4; CHST5; CST; HNK-1ST; SULFAL; SULFBL; SULFX3; STE; TPST2.
XX
XX OS Homo sapiens.
XX
XX JP2002085067-A.
XX
XX 26-MAR-2002.
XX
XX 07-SEP-2000; 2000JP-0272229.
XX 07-SEP-2000; 2000JP-0272229.
XX
XX (SAKA ) OTSUKA SEIYAKU KOGYO KK.
XX
XX WPI; 2002-378272/41.
XX
XX Determination of enzymes participating in sulphuric acid conjugation in
XX PT humans, useful for confirmation of safety of investigational drugs,
XX PT comprises using oligonucleotide probes
XX
XX Claim 4; Page 9; 13pp; Japanese.
XX
XX The invention relates to classification and quantitative determination of
XX CC enzymes participating in sulphuric acid conjugation comprising using

```

CC oligonucleotide probes hybridizing to the following regions: (a) 885-911
CC region of CHST1 gene; (b) 174-197 region of CHST3 gene; (c) 1003-1032
CC region of CHST4 gene; (d) 322-346 region of CHST5 gene; (e) 737-765
CC region of CST gene; (f) 703-732 region of HNK-1ST gene; (g) 299-325
CC region of SULTR2AI gene; (h) 358-382 region of SULTR2BI gene; (i) 554-582
CC region of SULTR3 gene; (j) 451-478 region of STB gene; and (k) 652-677
CC region of TPST2 gene. Also included are PCR primers for the above
CC genes, kits and methods for determination. The probes, primers and the
CC method are used in the determination of sulphuric acid conjugation for
CC confirmation of the safety of investigational drugs. The present
CC sequence is an oligonucleotide probe for one of the above listed genes.
XX

SO Sequence 30 BP; 5 A; 7 C; 7 G; 11 T; 0 other;

Query Match

Best Local Similarity 1.5%; Score 30; DB 24; Length 30;
Best Local Similarity 100.0%; Pred. No. 0.0096;
Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1184 CGCTGGCTCTTGGCCATGAAAGCTTTC 1213
DB 1 CGCTGGCTCTTGGCCATGAAAGCTTTC 30

RESULT 21

AAZ20793
ID AAZ20793 standard; DNA; 1926 BP.
XX
AC AAZ20793;
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 coding sequence.
XX
KM Glycosyl sulfotransferase: GST-3; detection; diagnosis; leukocyte homing;
KM selectin binding interaction; inflammation; lymphocyte homing; mouse;
KM secondary lymph organ; ss.
XX
OS Mus sp.
XX
PN WO9449018-A1.
XX
PD 30-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIT CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
XX
XX WPI; 1999-580442/49.
DR P-PSDB; AAY39919.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
XX polynucleotides
PS Claim 4; Fig 3; 59pp; English.
XX
CC This sequence encodes the mouse glycosyl sulfotransferase-3 (GST-3) of
CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC The sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions,
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.
XX
SQ Sequence 1926 BP; 426 A; 520 C; 484 G; 496 T; 0 other;

Query Match

1.3%; Score 27; DB 20; Length 1926;

Best Local Similarity 100.0%; Pred. No. 0.14;
Matches 27; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 494 GACATGACGCTCTTGATGACCTACATG 520
DB 714 GACATGACGCTCTTGATGACCTACATG 740

Search completed: February 22, 2003, 20:44:18
Job time : 697 secs

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 20:29:50 ; Search time 95 Seconds

(without alignments)
6559.655 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032 1 ggctcgagggccaggaatgacct.....ataagaaaaa 2032

Sequence: 1 ggctcgagggccaggaatgacct.....ataagaaaaa 2032

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 441362 seqs, 15338381 residues

Word size: 25

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Issued Patents -NA: *
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq: *
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq: *
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq: *
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq: *
5: /cgn2.6/ptodata/1/ina/PTUS.COMB.seq: *
6: /cgn2.6/ptodata/1/ina/Backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	2032	100.0	2032	4	US-09-045-284A-1
2	2032	100.0	2032	4	US-09-190-911-2
3	37	1.8	37	4	US-09-045-284A-7
4	37	1.8	37	4	US-09-190-911-7

ALIGNMENTS

RESULT 1
US-09-045-284A-1
; Sequence 1, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107051
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2032
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-045-284A-1

Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	GGCTCGAGGCGCAGATGCTCCAGTCTGGGGGAAAATGCTTCCTCATTTGCTTCCAG	60
DB	1	GGCTCGAGGCGCAGATGCTCCAGTCTGGGGGAAAATGCTTCCTCATTTGCTTCCAG	60
QY	61	CCACCTCAAGCAGTCTCCACCCCTGAGTCTCAGAGGTAAAGCTTTCTTCA	120
DB	61	CCACCTCAAGCAGTCTCCACCCCTGAGTCTCAGAGGTAAAGCTTTCTTCA	120
QY	121	CAGCTTCCTGGGAGGAGTCTTCTCAAGCCGCTTTCAGAGGTCTTCCACTTCAGCAC	180
DB	121	CAGCTTCCTGGGAGGAGTCTTCTCAAGCCGCTTTCAGAGGTCTTCCACTTCAGCAC	180
QY	181	AATGCTACTGCTTAAAAAATGAAGCTCTGCTTTCTGCTTTCCAGATGGCCATCTT	240
DB	181	AATGCTACTGCTTAAAAAATGAAGCTCTGCTTTCTGCTTTCCAGATGGCCATCTT	240
QY	241	GGCTCTATTTCCACATGTACAGCACAATCAGCTCCCTGCTTGAAGGACAGCC	300
DB	241	GGCTCTATTTCCACATGTACAGCACAATCAGCTCCCTGCTTGAAGGACAGCC	300
QY	301	CGAGCGCATGCAGTGTCTGTCTCTCTGCGGCTCTGCTCTTCTTTTGGGGCA	360
DB	301	CGAGCGCATGCAGTGTCTGTCTCTCTGCGGCTCTGCTCTTCTTTTGGGGCA	360
QY	361	GCTTTTGGGAGGACCCAGATGTTTCTTACCTGATGAGGCGCGCGGACGCTGGAT	420
DB	361	GCTTTTGGGAGGACCCAGATGTTTCTTACCTGATGAGGCGCGCGGACGCTGGAT	420
QY	421	GACCTTCAAGCAGACCGGCTGAGTCTGACATGGCTCTGGGGAATCATACGGGC	480
DB	421	GACCTTCAAGCAGACCGGCTGAGTCTGACATGGCTCTGGGGAATCATACGGGC	480
QY	481	CGCTCTCTTGGGAGCAGATGAGTCTTGTATGCTTACATGAGAACCTGCTCCCGAGACA	540
DB	481	CGCTCTCTTGGGAGCAGATGAGTCTTGTATGCTTACATGAGAACCTGCTCCCGAGACA	540
QY	541	GTCAGCCCTCTTTCAGTGGGAGAACAGCGGCGCTGTCTGACCTGCTCTGACAT	600
DB	541	GTCAGCCCTCTTTCAGTGGGAGAACAGCGGCGCTGTCTGACCTGCTCTGACAT	600
QY	601	CATCCCAAAATGAAATCATCCCGGGCTCAGTCAAGGCTCTCTGTCATCAACAGCC	660
DB	601	CATCCCAAAATGAAATCATCCCGGGCTCAGTCAAGGCTCTCTGTCATCAACAGCC	660
QY	661	CTTGAAGTGTGGAGAGGCTGCGCTCTACAGCAGCTGCTGTCAGAGAGGTGGC	720
DB	661	CTTGAAGTGTGGAGAGGCTGCGCTCTACAGCAGCTGCTGTCAGAGAGGTGGC	720
QY	721	CTTCTTCAACCTGCAAGTCTCTTACCGGCTGCTGAAGACCCCTCTCAACCTGATAT	780
DB	721	CTTCTTCAACCTGCAAGTCTCTTACCGGCTGCTGAAGACCCCTCTCAACCTGATAT	780
QY	781	CGTGACCTGGTCCGGGAGCCCGGGCGGTCTGCTCCGAGAGCAAAAGGGAGA	840
DB	781	CGTGACCTGGTCCGGGAGCCCGGGCGGTCTGCTCCGAGAGCAAAAGGGAGA	840
QY	841	TCTCATGATGACAGTGCATTTGATGGGACAGATGAGAGAACTCAAGAGAGAGA	900
DB	841	TCTCATGATGACAGTGCATTTGATGGGACAGATGAGAGAACTCAAGAGAGAGA	900
QY	901	CCAACTTACTATGATGATGAGTCACTGCTGCAAGCCAGCTGAGATCTCAAGACAT	960
DB	901	CCAACTTACTATGATGATGAGTCACTGCTGCAAGCCAGCTGAGATCTCAAGACAT	960
QY	961	CCAGTCTTGGCCAAAGGCGCTGCAAGAGCTTACTGCTTGGGCTTGAAGAGCTGGC	1020
DB	961	CCAGTCTTGGCCAAAGGCGCTGCAAGAGCTTACTGCTTGGGCTTGAAGAGCTGGC	1020

1021 TCGAGCCCTGTGGCCAGACTTCCGAGATGATGATTCGTGGATTTGGATTTGCC 1080
1021 TCGAGCCCTGTGGCCAGACTTCCGAGATGATGATTCGTGGATTTGGATTTGCC 1080
1081 CCAATCTTACAGCTGGGTGATTAACATCACCCGAGGCAAGGGATGGGTGACACGCTTT 1140
1081 CCAATCTTACAGCTGGGTGATTAACATCACCCGAGGCAAGGGATGGGTGACACGCTTT 1140
1141 CCACACAAATGCGAGGATGCTTATGCTCCAGGCTTTGGGCGCTTGTGGCCCTA 1200
1141 CCACACAAATGCGAGGATGCTTATGCTCCAGGCTTTGGGCGCTTGTGGCCCTA 1200
1201 TGAAGAGTTCTTCTGACTTCAAGAAAGCTTGGCGATGATGATTAATTTGCTGACCG 1260
1201 TGAAGAGTTCTTCTGACTTCAAGAAAGCTTGGCGATGATGATTAATTTGCTGACCG 1260
1261 CCAGCTGAGATTCAGACAGAGAGAGAAAGCTTGGCTGATGATTTCTTCTGCTCCGGAC 1320
1261 CCAGCTGAGATTCAGACAGAGAGAGAAAGCTTGGCTGATGATTTCTTCTGCTCCGGAC 1320
1321 TGTCCCTGAGCAATCCACTAAGAGGTTGAGAAGGCTTTGCTGCCACCTGCTGACGCC 1380
1321 TGTCCCTGAGCAATCCACTAAGAGGTTGAGAAGGCTTTGCTGCCACCTGCTGACGCC 1380
1381 TCAATCACTTCTCTGATGATGCTTCTGAGCTTGCCTACATCTGAGACCTTACTACATG 1440
1381 TCAATCACTTCTCTGATGATGCTTCTGAGCTTGCCTACATCTGAGACCTTACTACATG 1440
1441 TCTGTGGGTATACACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1441 TCTGTGGGTATACACACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1500
1501 GTGTCCATGCTTGTGTAGAAAACAGAGCTGGGGAACCTTATGTGAGACAGACATCCAC 1560
1501 GTGTCCATGCTTGTGTAGAAAACAGAGCTGGGGAACCTTATGTGAGACAGACATCCAC 1560
1561 CAGTGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1561 CAGTGAACAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
1621 AGACTTGTGGCTGAGAGGCTTATTAAGACAGACAGATGATGATGATGATGATGATGATGAT 1680
1621 AGACTTGTGGCTGAGAGGCTTATTAAGACAGACAGATGATGATGATGATGATGATGATGAT 1680
1681 ACCCTGCTGACAGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCT 1740
1681 ACCCTGCTGACAGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCT 1740
1741 ATTTTCCAGAGAGATGCAATTTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAG 1800
1741 ATTTTCCAGAGAGATGCAATTTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAGGCTTGTGAG 1800
1801 GTGGGAACAAAGTTGATGCTTATGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 1860
1801 GTGGGAACAAAGTTGATGCTTATGAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAGCT 1860
1861 AATATGAACAAATCTCTGACAAAAGAGCAAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 1920
1861 AATATGAACAAATCTCTGACAAAAGAGCAAGCTTGTGAGCTTGTGAGCTTGTGAGCTTGTGAG 1920
1921 TGAATTTGAATATCACTTCCCTCTGATTTTCCATCAATAGAGAGACTTTGACCTGTG 1980
1921 TGAATTTGAATATCACTTCCCTCTGATTTTCCATCAATAGAGAGACTTTGACCTGTG 1980
1981 AAGCTGCACTCTGATTAATTAATTTCCAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2032
1981 AAGCTGCACTCTGATTAATTAATTTCCAAATAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2032

RESULT 2
US-09-190-911-2
; Sequence 2, Application US/09190911

Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/09/190,911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045,284
EARLIER FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 2
TYPE: DNA
ORGANISM: H. sapiens
US-09-190-911-2
Query Match 100.0%; Score 2032; DB 4; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 GGTGAGGCGCAGAGATGCTCCAGTCTGGGGGAAATGCTTCTGATTTGCTTCCAG 60
1 GGTGAGGCGCAGAGATGCTCCAGTCTGGGGGAAATGCTTCTGATTTGCTTCCAG 60
61 CCCACTCAAGCAGATGCTCCACCCCTTGTAGTCTGAGCAGATTAAGCTGTACTTCA 120
61 CCCACTCAAGCAGATGCTCCACCCCTTGTAGTCTGAGCAGATTAAGCTGTACTTCA 120
121 CAGCTTCTCGGAGGAGAGTCTTCTCAAGCCCGTCTGCAAGGCTTCTCACTTACAG 180
121 CAGCTTCTCGGAGGAGAGTCTTCTCAAGCCCGTCTGCAAGGCTTCTCACTTACAG 180
181 AATGCTAGCTTAATAAATGAAGCTCTGCTGTTTCCAGTTCAGATGAGCATTT 240
181 AATGCTAGCTTAATAAATGAAGCTCTGCTGTTTCCAGTTCAGATGAGCATTT 240
241 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
241 GGTCTATTTCTTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 300
301 GAGGCGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
301 GAGGCGATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
361 GCTTTTGGGAGCAGCCAGATGTTTCTTACCTGATGAGCCGCTGAGCTGTGAT 420
361 GCTTTTGGGAGCAGCCAGATGTTTCTTACCTGATGAGCCGCTGAGCTGTGAT 420
421 GACCTTCAAGCAGAGCAGCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
421 GACCTTCAAGCAGAGCAGCAGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTGAT 480
481 CCGTCTTCTGTCGACATGAGGCTGTTGATGCTTACATGAGAACTGCTGCTCCGAGACA 540
481 CCGTCTTCTGTCGACATGAGGCTGTTGATGCTTACATGAGAACTGCTGCTCCGAGACA 540
541 GTTCAGCTCTTCTGAGTGGAGAGACAGCCGAGCTGTGTTGTGACCTGCTGTGAT 600
541 GTTCAGCTCTTCTGAGTGGAGAGACAGCCGAGCTGTGTTGTGACCTGCTGTGAT 600
601 GTTCAGCTCTTCTGAGTGGAGAGACAGCCGAGCTGTGTTGTGACCTGCTGTGAT 660
601 GTTCAGCTCTTCTGAGTGGAGAGACAGCCGAGCTGTGTTGTGACCTGCTGTGAT 660
661 CATTCCCAAGATGAAATATATCCCGGAGCTACAGAGCTTCTGTGATGATCAAGAGCC 720
661 CATTCCCAAGATGAAATATATCCCGGAGCTACAGAGCTTCTGTGATGATCAAGAGCC 720
721 CTCTTCAACCTGAGCTGCTTACCCGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
721 CTCTTCAACCTGAGCTGCTTACCCGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAG 780

Db	721	CTTCTTCAACCTGCAAGTCCCTCTACCCCGGTGCTGAAGAAGCCCTCCCTCAACTCATAT	780
Qy	781	CGTGCACCTGGTCCGGGAACCCCCGGGCGGTTCCTTCCCGAAGAACGACAAAGGAGA	840
Db	781	CGTGCACCTGGTCCGGGAACCCCCGGGCGGTTCCTTCCCGAAGAACGACAAAGGAGA	840
Qy	841	TCTCATGATTGACAGTGTGCATTGTGATGAGGGGACAGATGAGACAGAAACTCAAGAAGAGA	900
Db	841	TCTCATGATTGACAGTGTGCATTGTGATGAGGGGACAGATGAGACAGAAACTCAAGAAGAGA	900
Qy	901	CCAAAGCTACTATGTGATGTCAGAGTTCATCTGCCAAGCCAGCTGGAGATCTACAGACCAT	960
Db	901	CCAAAGCTACTATGTGATGTCAGAGTTCATCTGCCAAGCCAGCTGGAGATCTACAGACCAT	960
Qy	961	CCAGTCCCTGGCCAAAGCCCTGCAAGAACCTACCTGCTGTGCGGTATGAGACCTGGC	1020
Db	961	CCAGTCCCTGGCCAAAGCCCTGCAAGAACCTACCTGCTGTGCGGTATGAGACCTGGC	1020
Qy	1021	TCGAGCCCTGTGGCCCGACCTTCCGATGTATGAAITGTGGGATTGGAATTCYTGGC	1080
Db	1021	TCGAGCCCTGTGGCCCGACCTTCCGATGTATGAAITGTGGGATTGGAATTCYTGGC	1080
Qy	1081	CCATCTTCAGACCTGGGTCGATTAACATACCCGAGCAGGGCATGGGTCAGACAGCTTT	1140
Db	1081	CCATCTTCAGACCTGGGTCGATTAACATACCCGAGCAGGGCATGGGTCAGACAGCTTT	1140
Qy	1141	CCACACAAATGCCAGGGATCCCTTATGTCTCCAGGCTTGGCGGTGGTCTTGGCCTA	1200
Db	1141	CCACACAAATGCCAGGGATCCCTTATGTCTCCAGGCTTGGCGGTGGTCTTGGCCTA	1200
Qy	1201	TGAAAAGSTTCTCGACTTCAGAAAGCTGTGGCATGCCATGATTTCTGGGCTACCG	1260
Db	1201	TGAAAAGSTTCTCGACTTCAGAAAGCTGTGGCATGCCATGATTTCTGGGCTACCG	1260
Qy	1261	CCAGCTCAGATCTGAACAACAGACAGAAACCTGTGTGATGATCTGTCTACTACCTGGAC	1320
Db	1261	CCAGCTCAGATCTGAACAACAGACAGAAACCTGTGTGATGATCTGTCTACTACCTGGAC	1320
Qy	1321	TGTCCCTGAGCAAACTCCACTAAGAGGGTGTGAGAAAGCTTGTGTCGCCACTGGTGTACGC	1380
Db	1321	TGTCCCTGAGCAAACTCCACTAAGAGGGTGTGAGAAAGCTTGTGTCGCCACTGGTGTACGC	1380
Qy	1381	TCAGTCACTTTCGTGTAATGCTTCTGAGCTTGCCTACACATCTGACCTTACTACATG	1440
Db	1381	TCAGTCACTTTCGTGTAATGCTTCTGAGCTTGCCTACACATCTGACCTTACTACATG	1440
Qy	1441	TCTGTGGGTATCACACTGAGTGTGAGTGTGTGTCCACACCTGCTCAAGCAGAAAGCATTT	1500
Db	1441	TCTGTGGGTATCACACTGAGTGTGAGTGTGTGTCCACACCTGCTCAAGCAGAAAGCATTT	1500
Qy	1501	GTGTCCATCGTGTGTGTATAAAAAACAGACITGGGAAACCTTATGTAGACAGCACATCCAC	1560
Db	1501	GTGTCCATCGTGTGTGTATAAAAAACAGACITGGGAAACCTTATGTAGACAGCACATCCAC	1560
Qy	1561	CAGTGAACACAGGATATGTGCTTCTCTTCTTGTGATCTTCCGTGTGGGACGACTTCAG	1620
Db	1561	CAGTGAACACAGGATATGTGCTTCTCTTCTTGTGATCTTCCGTGTGGGACGACTTCAG	1620
Qy	1621	AGACTTGTGGCCTGAGGGCCTATTAAAGCAGCAGACAGTATCAGTGTGAAATTGATCATTA	1680
Db	1621	AGACTTGTGGCCTGAGGGCCTATTAAAGCAGCAGACAGTATCAGTGTGAAATTGATCATTA	1680
Qy	1681	ACGTCCCGTCCACATCTTGGCCCAATGGGAAATGGAATCTTTCACAAAGAGCTCACAGC	1740
Db	1681	ACGTCCCGTCCACATCTTGGCCCAATGGGAAATGGAATCTTTCACAAAGAGCTCACAGC	1740
Qy	1741	ATTTCACAGAGATTCCAAAATTCAGAGCCCTTGAGATTTCCAGTGTGATTCAGAGAGAA	1800
Db	1741	ATTTCACAGAGATTCCAAAATTCAGAGCCCTTGAGATTTCCAGTGTGATTCAGAGAGAA	1800
Qy	1801	GTGGGAACAGGTTGATGCTTATGTAGCTTGACCATCAGCTATCGGTAAATCAGA	1860
Db	1801	GTGGGAACAGGTTGATGCTTATGTAGCTTGACCATCAGCTATCGGTAAATCAGA	1860

```

Db      1801 GTGGGAACAAGSTTGATGCTACTTATTAGCTTGACCATCAGCTATACGTAATCAGAA 1860
QY      1861 AATATGAACAAATATCTCTGCAACAAAAGAGCAAGCTCTTAATTCACAGGTCGCTGGGC 1920
Db      1861 AATAAGAAACAAAATCTCTGCAACAAAAGAGCAAGCTCTTAATTCACAGGTCGCTGGGC 1920
QY      1921 TGCATTGGAATATACACTTCCCTCCCTGTCATTTCCCATCAGATAGAAAGACTTTGACCTGTG 1980
Db      1921 TGCATTGGAATATACACTTCCCTCCCTGTCATTTCCCATCAGATAGAAAGACTTTGACCTGTG 1980
QY      1981 AAGCTGCCATCTGTTAATCTATAAATTCCTCAATAATAGAAAAAATATATATATATATAT 2032
Db      1981 AAGCTGCCATCTGTTAATCTATAAATTCCTCAATAATAGAAAAAATATATATATATATAT 2032

RESULT 3
US-09-045-284A-7
; Sequence 7, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerlich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107051
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 7
; LENGTH: 37
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-045-284A-7

Query Match          1.8%: Score 37; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 4,3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      884 AACTCAAGAAAGAGAGACCACCACTACTATGTGATGC 920
Db      1 AACTCAAGAAAGAGAGACCACCACTACTATGTGATGC 37

RESULT 4
US-09-190-911-7
; Sequence 7, Application US/09190911
; Patent No. 6365365
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Tagemann, Kirsten
; APPLICANT: Hemmerlich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CIP
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 7
; LENGTH: 37
; TYPE: DNA
; ORGANISM: H. sapiens
US-09-190-911-7

Query Match          1.8%: Score 37; DB 4; Length 37;
Best Local Similarity 100.0%; Pred. No. 4,3e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      884 AACTCAAGAAAGAGAGACCACCACTACTATGTGATGC 920

```

Db 1 AAAGTCAAGAGAGAGACCAACCTACTATGTGATGC 37

Search completed: February 22, 2003, 22:53:37
Job time : 96 secs

Mon Feb 24 10:56:26 2003

us-09-816-825-1.olig25.rnpb

Page 1

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 22, 2003, 22:24:31 ; Search time 166 Seconds

(without alignments)
6875.008 Million cell updates/sec

Title: US-09-816-825-1

Perfect score: 2032
Sequence: 1 ggcgcagagcagagatgcct.....ataaagaaaaaaaaa 2032

Scoring table: OLIGO-MNC
Gapop 60.0 , Gapext 60.0

Searched: 442118 seqs, 280819700 residues

Word size : 25

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Published_Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCR_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCRUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2032	100.0	2032	9	US-10-007-262-2
2	2032	100.0	2032	10	US-09-816-825-1
3	505	24.9	505	10	US-09-998-598-2595
4	505	24.9	517	10	US-09-919-580-194
5	286	14.1	389	10	US-09-919-580-264
6	59	2.9	2544	10	US-09-927-602-1
7	59	2.9	48436	10	US-09-927-602-38
8	37	1.8	37	9	US-10-007-262-7
9	37	1.8	37	10	US-09-816-825-7

ALIGNMENTS

RESULT 1
US-10-007-262-2
; Sequence 2, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:

APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OR INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2032
TYPE: DNA
ORGANISM: H. sapiens
US-10-007-262-2

Query Match 100.0%; Score 2032; DB 9; Length 2032;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2032; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCCTGAGGCGCAGATGCTCCAGCTGAGGGAATGCTTCCTGCTCTCCAG 60
DB 1 GGCCTGAGGCGCAGATGCTCCAGCTGAGGGAATGCTTCCTGCTCTCCAG 60
QY 61 CCCACTCAAGACAGCTCCACCCCTTGAGTCCAGAGTGAAGCTGTTACTTCA 120
DB 61 CCCACTCAAGACAGCTCCACCCCTTGAGTCCAGAGTGAAGCTGTTACTTCA 120
QY 121 CAGCTTCTGGAGGAGGAGTCTTCTCAAGCCGCTTCAGAGTCTCCACTTCA 180
DB 121 CAGCTTCTGGAGGAGGAGTCTTCTCAAGCCGCTTCAGAGTCTCCACTTCA 180
QY 181 AATCTACTGCTCTCAATGAAGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 240
DB 181 AATCTACTGCTCTCAATGAAGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 240
QY 241 GGCCTCTATCTTCCATGTAACGCAACATCATGCTCTGCTCTGCTCTGCT 300
DB 241 GGCCTCTATCTTCCATGTAACGCAACATCATGCTCTGCTCTGCTCTGCT 300
QY 301 CGAGCGCATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CGAGCGCATGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 GCTTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 361 GCTTTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 GACCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 421 GACCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GCTTCTTGTGCGACATGAGCGCTTGTGATGCTTACATGGAACCTGCTCCG 540
DB 481 GCTTCTTGTGCGACATGAGCGCTTGTGATGCTTACATGGAACCTGCTCCG 540
QY 541 GTCAGGCTCTTTCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
DB 541 GTCAGGCTCTTTCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
QY 601 CATCCCAAGAGTGAATCAATCCCGGGGCTCACTGAGGCTCCTGAGCTCAACG 660
DB 601 CATCCCAAGAGTGAATCAATCCCGGGGCTCACTGAGGCTCCTGAGCTCAACG 660
QY 661 CTTTGAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
DB 661 CTTTGAAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 CTTCTTCAACCTGAGTCCCTTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCT 780
DB 721 CTTCTTCAACCTGAGTCCCTTCAACGCTGCTGCTGCTGCTGCTGCTGCTGCT 780

Db 481 CGCTCTCTGTGCGACATGAGCGCTTTGATGCTTACATGAAACCGTCCCCGGAGACA 540
QY 541 GTCCAGCCCTCTTTCAGTGGGAGAAACGCCGGCCCTGTGTTCTGCACCTGCTGTGACAT 600
Db 541 GTCCAGCCCTCTTTCAGTGGGAGAAACGCCGGCCCTGTGTTCTGCACCTGCTGTGACAT 600
QY 601 CATCCACAGATGAATCATCCCCGGGCTACTGCAGGCTCCGTGAGTCAACAGCC 660
Db 601 CATCCACAGATGAATCATCCCCGGGCTACTGCAGGCTCCGTGAGTCAACAGCC 660
QY 661 GTTGAAGTGTGAGAAAGGCTGCCCTCTACAGCCAGTGTGCTCAAGAGAGTGC 720
Db 661 GTTGAAGTGTGAGAAAGGCTGCCCTCTACAGCCAGTGTGCTCAAGAGAGTGC 720
QY 721 GTTCTCAACCTGAGTCCCTCTACCCGGTGTGAAAGACCCCTCCCTCAACCTGCATAT 780
Db 721 GTTCTCAACCTGAGTCCCTCTACCCGGTGTGAAAGACCCCTCCCTCAACCTGCATAT 780
QY 781 CGTGCACCTGTGTCGGGACCCCGCGGTGTTCCGTTCCCGAGAACCCACAAAGGAGA 840
Db 781 CGTGCACCTGTGTCGGGACCCCGCGGTGTTCCGTTCCCGAGAACCCACAAAGGAGA 840
QY 841 TGTCAATGATGACATGCGATTTGTATGGGAGCATGAGCAAGAACTCAAGAGAGAGA 900
Db 841 TGTCAATGATGACATGCGATTTGTATGGGAGCATGAGCAAGAACTCAAGAGAGAGA 900
QY 901 CCACCCCTACTATGTATGATGACAGTATGCAAGCCAGCTGAGATCTACAAAGCAT 960
Db 901 CCACCCCTACTATGTATGATGACAGTATGCAAGCCAGCTGAGATCTACAAAGCAT 960
QY 961 CCAGTCTCTGGCCCAAGGCCCTGACAGAAACCTACTGCTGTGGGCTATGAGACCTGGC 1020
Db 961 CCAGTCTCTGGCCCAAGGCCCTGACAGAAACCTACTGCTGTGGGCTATGAGACCTGGC 1020
QY 1021 TCGAGCCCTGTGTCGGGACCTCCCGAATGTATGATTCGTGGGATTTGGAATTTGCC 1080
Db 1021 TCGAGCCCTGTGTCGGGACCTCCCGAATGTATGATTCGTGGGATTTGGAATTTGCC 1080
QY 1081 CCATTTTCAAGACCTGGGTGCATACATACCCGAGCAAGGCGCATGGGTGACACGCTTT 1140
Db 1081 CCATTTTCAAGACCTGGGTGCATACATACCCGAGCAAGGCGCATGGGTGACACGCTTT 1140
QY 1141 CCACACAATGGCAGGAGTGCCTTAATGTCTCCAGGCTTGGGCTGTGCTTGGCCTA 1200
Db 1141 CCACACAATGGCAGGAGTGCCTTAATGTCTCCAGGCTTGGGCTGTGCTTGGCCTA 1200
QY 1201 TGAAGAAGTTCTCGACTTCAGAAAGCTGTGGCGATGCATGAATTTGCTGGGCTACCG 1260
Db 1201 TGAAGAAGTTCTCGACTTCAGAAAGCTGTGGCGATGCATGAATTTGCTGGGCTACCG 1260
QY 1261 CCAGTCAATGATGACAAAGAAAGAAACCTGTGTGATGCTTCTGTCTACCTGAGC 1320
Db 1261 CCAGTCAATGATGACAAAGAAAGAAACCTGTGTGATGCTTCTGTCTACCTGAGC 1320
QY 1321 TGTCCCTGAGCAAAATCCATGAAGAGGTTGAGAAAGGCTTGTCTGCACCTGCTGACCC 1380
Db 1321 TGTCCCTGAGCAAAATCCATGAAGAGGTTGAGAAAGGCTTGTCTGCACCTGCTGACCC 1380
QY 1381 TCAATGATTTCTCTGAAATGCTTCTGAGACCTTGCTACATGCTGAGCCTTAACATCATG 1440
Db 1381 TCAATGATTTCTCTGAAATGCTTCTGAGACCTTGCTACATGCTGAGCCTTAACATCATG 1440
QY 1441 TCTGTGGTATACACACTAGTGTGAGTGTGTCCACAGTCTCAAGAGAGAGCTTT 1500
Db 1441 TCTGTGGTATACACACTAGTGTGAGTGTGTCCACAGTCTCAAGAGAGAGCTTT 1500
QY 1501 GTGTCCATGCTTTGTCTAGAAAACAGACTGGGAACTTATGTAGAGACACATCCAC 1560
Db 1501 GTGTCCATGCTTTGTCTAGAAAACAGACTGGGAACTTATGTAGAGACACATCCAC 1560
QY 1561 CAGTGAACAGAGGTTATGCTCTTCTTCTTCTGATGATCTTCTGCTGAGAGACTTCAG 1620
Db 1561 CAGTGAACAGAGGTTATGCTCTTCTTCTTCTGATGATCTTCTGCTGAGAGACTTCAG 1620

QY 1621 AGACTTTGTGCGCTTGAGAGCCTATTAAACAGACAGATATCACTGGAATTTGATCATAA 1680
Db 1621 AGACTTTGTGCGCTTGAGAGCCTATTAAACAGACAGATATCACTGGAATTTGATCATAA 1680
QY 1681 ACCTCCCTGCTACATCTTGCCTCAATGGGGAATGATCTTCCACAAAGAGCTACAGC 1740
Db 1681 ACCTCCCTGCTACATCTTGCCTCAATGGGGAATGATCTTCCACAAAGAGCTACAGC 1740
QY 1741 ATTTCCACAGAGATGCAAAATTTCTGAGCCCTTGAGAGTTCCTGATTTCAAGAGAGAA 1800
Db 1741 ATTTCCACAGAGATGCAAAATTTCTGAGCCCTTGAGAGTTCCTGATTTCAAGAGAGAA 1800
QY 1801 GTGGGAACAGGTTGAGTGGTCTACTTATGAGCTTGACATACAGCTATCGGTATATAGA 1860
Db 1801 GTGGGAACAGGTTGAGTGGTCTACTTATGAGCTTGACATACAGCTATCGGTATATAGA 1860
QY 1861 AATATGAACAAATCTCTGACACAAAGAGCAAGCTTTAAGTTCACAGGAGTGCCTGGC 1920
Db 1861 AATATGAACAAATCTCTGACACAAAGAGCAAGCTTTAAGTTCACAGGAGTGCCTGGC 1920
QY 1921 TGCATTTGAATATACATTTCCCTCTGCATTTTCCATCACAATAGAACATTTGACCTGTG 1980
Db 1921 TGCATTTGAATATACATTTCCCTCTGCATTTTCCATCACAATAGAACATTTGACCTGTG 1980
QY 1981 AAGCTGCATCTGTTAATACTAAATTCCTCAATTAAGAAAAA 2032
Db 1981 AAGCTGCATCTGTTAATACTAAATTCCTCAATTAAGAAAAA 2032

RESULT 3
US-09-998-2595
: Sequence 2595, Application US/0998598
: Patent No. US20020150922A1
: GENERAL INFORMATION:
: APPLICANT: Stolk, John A.
: APPLICANT: Xu, Jianshun
: APPLICANT: Chenault, Ruth A.
: APPLICANT: Meagher, Madeleine Joy
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
: FILE REFERENCE: 210121.561
: CURRENT APPLICATION NUMBER: US/09/998, 598
: CURRENT FILING DATE: 2001-11-16
: NUMBER OF SEQ ID NOS: 2606
: SOFTWARE: Corixa Invention Disclosure Database
: SEQ ID NO 2595
: LENGTH: 505
: TYPE: DNA
: ORGANISM: Homo sapiens
US-09-998-598-2595

Query Match 24.9%; Score 505; DB 10; Length 505;
Best Local Similarity 100.0%; Pred. No. 5, le-205;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 GGGCTTTGTGCGACCTGTGAGCTTCACTGCTGTAATGCTTGTGAGCCTTGC 1414
Db 1 GGGCTTTGTGCGACCTGTGAGCTTCACTGCTGTAATGCTTGTGAGCCTTGC 60
QY 1415 CTACATCTGAGGCTTAACATGATGCTGTGGGATACAGAGAGTGTGCTGTC 1474
Db 61 CTACATCTGAGGCTTAACATGATGCTGTGGGATACAGAGAGTGTGCTGTC 120
QY 1475 ACAGTCTCAAGAGAGAGCTTTTGTGTCATGCTTGTCTAGAAAACAGATGGG 1534
Db 121 ACAGTCTCAAGAGAGAGAGCTTTTGTGTCATGCTTGTCTAGAAAACAGATGGG 180
QY 1535 AACCTTGTGAGAGAGACATCCACAGAGTGAACAGAGATGCTTCTTTTCT 1594
Db 181 AACCTTGTGAGAGAGACATCCACAGAGTGAACAGAGATGCTTCTTTTCT 240
QY 1595 GATCTTCTGCTGTGGGAGACTTTCAGAGACTTGTGCTGAGAGGCTTATTAAGCAGAC 1654

```

Db 241 GATTTCTCTCTGGGAGACTTCAAGAGATTTGTGGCTGGAGGCTATTAAAGCAGAC 300
QY 1655 ACAGTATCAGTGAATGATCATAAACCTCCCTGTCACATCTTGGCCCAATGGGGAATG 1714
Db 301 ACAGTATCAGTGAATGATCATAAACCTCCCTGTCACATCTTGGCCCAATGGGGAATG 360
QY 1715 GATCTTCACCAAGAGCTCACAGCATTTTCACAGAGATGCAATTTGAGCCCTTGG 1774
Db 361 GATCTTCACCAAGAGCTCACAGCATTTTCACAGAGATGCAATTTGAGCCCTTGG 420
QY 1775 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAAGTGTGATGCTTACTTATGAGCTT 1834
Db 421 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAAGTGTGATGCTTACTTATGAGCTT 480
QY 1835 GACCATCACAGCTATGCTATGAG 1859
Db 481 GACCATCACAGCTATGCTATGAG 505

RESULT 4
US-09-919-580-194
; Sequence 194, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 194
; LENGTH: 517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-194

Query Match          24.9%; Score 505; DB 10; Length 517;
Best Local Similarity 100.0%; Pred. No. 5.1e-205;
Matches 505; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1355 GCGTTTGCTGCCACCTGGGTGTACAGCTCAGTCACTTTCTGTGATGCTTGTGAGCCTTGC 1414
Db 13 GCGTTTGCTGCCACCTGGGTGTACAGCTCAGTCACTTTCTGTGATGCTTGTGAGCCTTGC 72
QY 1415 CTACATCTGTGAGCTTAATACATGCTGTGGGTATACACAGTGAATGATGCTGTC 1474
Db 73 CTACATCTGTGAGCTTAATACATGCTGTGGGTATACACAGTGAATGATGCTGTC 132
QY 1475 ACAAGTGTCTAGCAGAGAGACTTTGTGTGCATGCTTGTGTAGAAAACAGACTGGGG 1534
Db 133 ACAAGTGTCTAGCAGAGAGACTTTGTGTGCATGCTTGTGTAGAAAACAGACTGGGG 192
QY 1535 AACCTTGTGAGAGAGACATCCACAGAGAGAGAGATGCTTGTGTGCTTCTTCTTCTT 1594
Db 193 AACCTTGTGAGAGAGACATCCACAGAGAGAGAGATGCTTGTGTGCTTCTTCTTCTT 252
QY 1595 GATCTTCCTGTCTGGGAGACTTCAGAGACTTTGTGGCTGGAGGCTTATTAAGCAGAC 1654
Db 253 GATCTTCCTGTCTGGGAGACTTCAGAGACTTTGTGGCTGGAGGCTTATTAAGCAGAC 312
QY 1655 ACAGTATCAGTGAATGATCATAAACCTCCCTGTCACATCTTGGCCCAATGGGGAATG 1714
Db 313 ACAGTATCAGTGAATGATCATAAACCTCCCTGTCACATCTTGGCCCAATGGGGAATG 372
QY 1715 GATCTTCACCAAGAGCTCACAGCATTTTCACAGAGATGCAATTTGAGCCCTTGG 1774
Db 373 GATCTTCACCAAGAGCTCACAGCATTTTCACAGAGATGCAATTTGAGCCCTTGG 432
```

```

QY 1775 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAAGTGTGATGCTTACTTATGAGCTT 1834
Db 433 AGTTCACAGTGAATCAAGAGAGAGTGGAAACAAGTGTGATGCTTACTTATGAGCTT 492
QY 1835 GACCATCACAGCTATGCTATGAG 1859
Db 493 GACCATCACAGCTATGCTATGAG 517

RESULT 5
US-09-919-580-264
; Sequence 264, Application US/09919580
; Patent No. US20020110832A1
; GENERAL INFORMATION:
; APPLICANT: Pyle, Ruth
; APPLICANT: Xu, Jiangchun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.552
; CURRENT APPLICATION NUMBER: US/09/919,580
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 934
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264
; LENGTH: 389
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-919-580-264

Query Match          14.1%; Score 286; DB 10; Length 389;
Best Local Similarity 100.0%; Pred. No. 5e-112;
Matches 286; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1732 CTCACAGCATTTTCCACAGAGATCCAAATTCGAGCCCTTGGAGTCCCAAGTGAATCA 1791
Db 13 CTCACAGCATTTTCCACAGAGATCCAAATTCGAGCCCTTGGAGTCCCAAGTGAATCA 72
QY 1792 AGGAAGAGAGTGGGAACAAGTGTGATGCTTACTTATGAGCTTGCATCACAGTATGC 1851
Db 73 AGGAAGAGAGTGGGAACAAGTGTGATGCTTACTTATGAGCTTGCATCACAGTATGC 132
QY 1852 GATATCAGAAATATGAACAAATCTGTGCACAAAGAGCAAGCTTAAAGTGCAGAGG 1911
Db 133 GATATCAGAAATATGAACAAATCTGTGCACAAAGAGCAAGCTTAAAGTGCAGAGG 192
QY 1912 TCCCTGGGCTGATTTGAATATCACTTCCCTGTGCATTTTCCATACATGAGACTT 1971
Db 193 TCCCTGGGCTGATTTGAATATCACTTCCCTGTGCATTTTCCATACATGAGACTT 252
QY 1972 TGACCTGTGAGCTGCATCTGTATATATAATTAATCCCAATTAAG 2017
Db 253 TGACCTGTGAGCTGCATCTGTATATATAATTAATCCCAATTAAG 298

RESULT 6
US-09-927-602-1
; Sequence 1, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2544
```

TYPE: DNA
ORGANISM: Homo Sapien
FEATURE:
NAME/KEY: CDS
LOCATION: (693)...(1877)
US-09-927-602-1

Query Match 2.9%; Score 59; DB 10; Length 2544;
Best Local Similarity 100.0%; Pred. No. 1.2e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGCGCTGCCCTCTACAGCCAGCTGTCACAGAGTGGCTTCTTCAACCTGCAG 736
DB 1180 AGCGCTGCCCTCTACAGCCAGCTGTCACAGAGTGGCTTCTTCAACCTGCAG 1238

RESULT 7
US-09-927-602-38
Sequence 38, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 38
LENGTH: 48436
TYPE: DNA
ORGANISM: Homo Sapien
US-09-927-602-38

Query Match 2.9%; Score 59; DB 10; Length 48436;
Best Local Similarity 100.0%; Pred. No. 1.1e-15;
Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 678 AGCGCTGCCCTCTACAGCCAGCTGTCACAGAGTGGCTTCTTCAACCTGCAG 736
DB 48024 AGCGCTGCCCTCTACAGCCAGCTGTCACAGAGTGGCTTCTTCAACCTGCAG 48082

RESULT 8
US-10-007-262-7
Sequence 7, Application US/10007262
Patent No. US20020164748A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 7
LENGTH: 37
TYPE: DNA
ORGANISM: H. sapiens
US-10-007-262-7

Query Match 1.8%; Score 37; DB 9; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTAAGAAGAGACCAACCCCTACTANGTATGTC 920
DB 1 AAACCTAAGAAGAGAGGACCAACCCCTACTANGTATGTC 37

RESULT 9
US-09-816-825-7
Sequence 7, Application US/09816825
Patent No. US20010051370A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 37
TYPE: DNA
ORGANISM: Homo sapiens
US-09-816-825-7

Query Match 1.8%; Score 37; DB 10; Length 37;
Best Local Similarity 100.0%; Pred. No. 2.8e-06;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 AAACCTAAGAAGAGACCAACCCCTACTANGTATGTC 920
DB 1 AAACCTAAGAAGAGAGGACCAACCCCTACTANGTATGTC 37

Search completed: February 23, 2003, 00:13:05
Job time : 270 secs

This page Blank (usps)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:22:21; Search time 35 Seconds
(without alignments)
2272.405 Million cell updates/sec

Title: US-09-816-825-2
Sequence: 1 MLPRKMKLLFLVSOMAIL.....EORNLDDLSTWTVPQIH 386

Scoring table: OLIGO
Gapop 60.0, Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 15

Total number of hits satisfying chosen parameters: 6

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

- SPTREMBL_21.*
- 1: sp-archaea.*
- 2: sp-bacteria.*
- 3: sp-fungi.*
- 4: sp-human.*
- 5: sp-invertebrate.*
- 6: sp-mammal.*
- 7: sp-mhc.*
- 8: sp-organelle.*
- 9: sp-phage.*
- 10: sp-plant.*
- 11: sp-prodent.*
- 12: sp-virus.*
- 13: sp-vertebrate.*
- 14: sp-unclassified.*
- 15: sp-virus.*
- 16: sp-bacteriap.*
- 17: sp-archaeap.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	386	100.0	386	4 Q9Y5R3	Q9Y5R3 homo sapien
2	39	10.1	388	11 Q9R1I1	Q9R1I1 mus musculi
3	39	10.1	388	11 Q9WJES	Q9WJES mus musculi
4	19	4.9	395	11 Q9GZX3	Q9GZX3 homo sapien
5	19	4.9	395	11 Q9QUP4	Q9QUP4 mus musculi
6	19	4.9	411	4 Q9GZS9	Q9GZS9 homo sapien

ALIGNMENTS

RESULT 1
Q9Y5R3 PRELIMINARY; PRT; 386 AA.
AC Q9Y5R3;
DT 01-NOV-1999 (TREMBLrel. 12, Created)

DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE	N-acetylglucosamine 6-O-sulfotransferase (U-selectin ligand
DE	sulfotransferase Gsr-3).
GN	GSR3.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=TONSIL;
RA	Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R.,
RA	Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.;
RT	"Sulfotransferases of two specificities function in the reconstitution
RT	of high endothelial cell ligands for U-selectin.";
RL	J. Cell Biol. 145:899-910(1999).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	TISSUE=TONSIL;
RA	Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
RA	Rosen S.D.;
RT	"Chromosomal localization and Genomic Organization for the
RT	Galactose/N-Acetylgalactosamine/N-Acetylglucosamine 6-O-
RT	Sulfotransferase Gene Family.";
RL	Glycobiology 0:0-0(2001).
RN	[3]
RP	SEQUENCE FROM N.A.
RC	MEDLINE=21332592; PubMed=11439191;
RA	Yeh J.-C., Hiraoka N., Petryniak B., Nakayama J., Ellies L.G.,
RA	Rabuka D., Hindsaul O., March J.D., Lowe J.B., Fakuda M.;
RT	"Novel sulfated lymphocyte homing receptors and their control by a
RT	corel extension betail,3-N-acetylglucosaminyltransferase.";
RL	Cell 105:957-969(2001).
DR	EMBL; AF131235; AAD33015.1; -
DR	EMBL; AF280088; AAG48246.1; -
DR	EMBL; AF149783; AAK48417.1; -
DR	InterPro; IPR001092; HELIX_BASIC.
DR	PROSITE; PS00036; HELIX_LOOP_HELIX; UNKNOWN_1.
KW	Lectin; Selectin; Transferase.
SQ	SEQUENCE 386 AA; 45133 MW; 0C3BB4022417143A CRC64;
QY	Query Match 100.0%; Score 386; DB 4; Length 386;
QY	Best Local Similarity 100.0%; Pred. No. 0;
QY	Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MLPRKMKLLFLVSOMAILALFFHMTSHNTSSLSMKRAQPERMHVLYSSWRSGSSFVGQ 60
DB	1 MLPRKMKLLFLVSOMAILALFFHMTSHNTSSLSMKRAQPERMHVLYSSWRSGSSFVGQ 60
QY	61 LFGQHPDYFLMEPAWVHWMTFKOSTAMLMANVRDLIRAVFLCDMSVFADYMEGRRQ 120
DB	61 LFGQHPDYFLMEPAWVHWMTFKOSTAMLMANVRDLIRAVFLCDMSVFADYMEGRRQ 120
QY	121 SSIFOWENSRAICGAPACDIIPODEIIPRAHCRILCSQOPEVEYKCRSYSHVLEKVR 180
DB	121 SSIFOWENSRAICGAPACDIIPODEIIPRAHCRILCSQOPEVEYKCRSYSHVLEKVR 180
QY	181 FPNQSLPLKDPSSLNHYHLVDRPAVFRSEERKGDIMDSRIYVMGQHEQKLEKED 240
DB	181 FPNQSLPLKDPSSLNHYHLVDRPAVFRSEERKGDIMDSRIYVMGQHEQKLEKED 240
QY	241 QPYVVMQVTCOSLEIKTIQSLPKALOERYILVRYEDLARAPAVQTSRYMEFGEFLP 300
DB	241 QPYVVMQVTCOSLEIKTIQSLPKALOERYILVRYEDLARAPAVQTSRYMEFGEFLP 300
QY	301 HLQTVNHNITRGKMGDAFTNARDALNVSQAWMSLPEYKYSRLQKACGDANMLLGYR 360
DB	301 HLQTVNHNITRGKMGDAFTNARDALNVSQAWMSLPEYKYSRLQKACGDANMLLGYR 360
QY	361 HVRSQDQRLNLDLSTWTVPQIH 386

Db 361 HVSEGEORNLIDLTSTWTFEQIH 386

RESULT 2

Q9RI11 ID PRELIMINARY; PRT; 388 AA.

AC Q9RI11; 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)

DE L-selectin ligand sulfotransferase.

GN CHST4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RE SEQUENCE FROM N.A.

RA MEDLINE-9361934; PubMed-10435561;

RA Hirata N., Petrylak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,

RA Iwata D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.;

RT "A novel, high endothelial venule-specific sulfotransferase expresses

RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.";

DR EMBL: AF109155; AAD45579.1; -

DR MGD; MGI:1349479; Chst4.

KW Lectin; Selectin; Transferase.

SO SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884AEE CRC64;

Query Match 10.1%; Score 39; DB 11; Length 388;

Best Local Similarity 100.0%; Pred. No. 1.4e-34;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSWRSGSFVQGLFGQHPDVFLMEPAHWMTF 82

Db 43 HVLVLSWRSGSFVQGLFGQHPDVFLMEPAHWMTF 81

RESULT 3

Q9WUE5 ID PRELIMINARY; PRT; 388 AA.

AC Q9WUE5; 01-NOV-1999 (Tremblrel. 12, Created)

DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE N-acetylglucosamine 6-O-sulfotransferase (Carbohydrate (Chondroitin

DE 6/Retan) sulfotransferase 4).

GN CHST4.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RE SEQUENCE FROM N.A.

RA STRAIN-C57BL/6; TISSUE-TONSIL;

RA MEDLINE-99264336; PubMed-10330415;

RA Bistrup A., Bhakta S., Lee J.-K., Below Y.Y., Gunn M.D., Zuo F.-R.,

RA Huang C.-C., Kannagi R., Rosen S.D., Hemmerich S.;

RT "Sulfotransferases of two specificities function in the reconstitution

RT of high endothelial cell ligands for L-selectin.";

RL J. Cell Biol. 145:899-910(1999).

RN [2]

RE SEQUENCE FROM N.A.

RA STRAIN-C57BL/6J; TISSUE-TONGUE;

RA MEDLINE-21085660; PubMed-11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,

RA Aizawa K., Iizawa M., Nishi K., Kiyosawa H., Konno S., Yamahata I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Cassavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Marzarelli J., Kommaerts P.,
RA Nordone P., Ring B., Rutgwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuki S.,
RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL; AF131236; AAD33016.1; -

DR MGD; AK009113; BAB26078.1; -

DR MGD; MGI:1349479; Chst4.

KW Transferase.

SO SEQUENCE 388 AA; 44694 MW; D1E9D796DF8574D CRC64;

Query Match

Best Local Similarity 10.1%; Score 39; DB 11; Length 388;

Matches 39; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSWRSGSFVQGLFGQHPDVFLMEPAHWMTF 82

Db 43 HVLVLSWRSGSFVQGLFGQHPDVFLMEPAHWMTF 81

RESULT 4

Q9GZX3 ID PRELIMINARY; PRT; 395 AA.

AC Q9GZX3; 01-MAR-2001 (Tremblrel. 16, Created)

DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)

DE N-acetylglucosamine 6-O-sulfotransferase GSN-4beta (Corneal N-

DE acetylglucosamine-6-O-sulfotransferase).

GN GSI4BETA OR CHST6.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RE SEQUENCE FROM N.A.

RA Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,

RA Rosen S.D.;

RT "Chromosomal localization and Genomic Organization for the

RT Galactose/N-Acetylglucosamine/N-Acetylglucosamine 6-O-

RT Sulfotransferase Gene Family.";

RL Glycobiology 0:0-0(2001).

RN [2]

RE SEQUENCE FROM N.A.

RA MEDLINE-20472330; PubMed-11017086;

RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Fujiwara T.,

RA Nakamura T., Dota A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S.,

RA Ozaki K., Kinoshita S., Shimomura Y., Tanigami A., Fukuda M.N.;

RT "Macular corneal dystrophy type I and type II are caused by distinct

RT mutations in a new sulphotransferase gene.";

RL Nat. Genet. 26:237-241(2000).

DR EMBL; AF280086; AAG48244.1; -

DR EMBL; AF219990; AAG26325.1; -

DR EMBL; AF219991; AAG26327.1; -

KW Transferase.

SO SEQUENCE 395 AA; 44098 MW; 433CA60248A48F7 CRC64;

Query Match

Best Local Similarity 4.9%; Score 19; DB 4; Length 395;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSWRSGSFVQGLF 62

Db 42 HVLVLSWRSGSFVQGLF 60


```

RESULT 5
Q9GZP4 PRELIMINARY; PRT; 395 AA.
AC Q9GZP4;
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase.
GN CHST5 OR I-GLCNAC-6-ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
RT Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF176841; AAD56003.1; -
DR EMBL; AF176840; AAD56002.1; -
DR MGD; MGI:1931825; Chst5.
KW SEQUENCE 395 AA; 44537 MW; 3FD71EA3ED383BE CRC64;

Query Match 4.9%; Score 19; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 3.5e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVLSNRSGSFEVQQLF 62
Db 42 HVLVLSNRSGSFEVQQLF 60

RESULT 6
Q9GZS9 PRELIMINARY; PRT; 411 AA.
AC Q9GZS9; Q9BY3;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Intestinal GlcNAc-6-sulfotransferase (Intestinal N-
DE acetylglucosamine-6-O-sulfotransferase) (N-acetylglucosamine
DE 6-O-sulfotransferase).
GN CHST5 OR I-GLCNAC-6-ST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20472330; PubMed=11017086;
RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K.,
RA Nakamura T., Dora A., Kawasaki S., Inoue Y., Maeda N., Yamamoto S.,
RA Fujiwara T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A.,
RA Fukuda M.N.;
RT "Macular corneal dystrophy type I and type II are caused by distinct
RT mutations in a new sulphotransferase gene.";
RL Nat. Genet. 26:237-241(2000).
RN [2]
RP SEQUENCE OF 22-411 FROM N.A.
RC TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
RT Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF246718; AAG28023.1; -
DR EMBL; AF219991; AAG26326.1; -
DR EMBL; AF176839; AAD56001.1; -
DR EMBL; AF176838; AAD56000.1; -

```

```

KW Transferase.
SQ SEQUENCE 411 AA; 46160 MW; 97642D54BE926E06 CRC64;

Query Match 4.9%; Score 19; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 3.6e-12;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSYSHVYKVRFFNLQ 185
Db 186 ACRSYSHVYKVRFFNLQ 204

```

Search completed: February 20, 2003, 13:25:37
Job time : 36 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:23:11 : Search time 18 Seconds
(without alignments)
630,958 Million cell updates/sec

Title: US-09-816-825-2

Perfect score: 386
Sequence: 1 MLPRKKMLLFLVSQMAIL.....EQRLLDLSTWTVEQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 15

Total number of hits satisfying chosen parameters: 2

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCNUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfilltest.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	386	100.0	386	4 US-09-045-284A-2	Sequence 2, Appli
2	386	100.0	386	4 US-09-190-911-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1

US-09-045-284A-2

Sequence 2, Application US/09045284A
Patent No. 6265192
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107051
CURRENT APPLICATION NUMBER: US/09/045, 284A
CURRENT FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 100.0%; Score 386; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKKMLLFLVSQMAILALFFHMYSHNLSLSKRAQPERMVLVLSWRSSSVGQ	60
DB	1	MLPRKKMLLFLVSQMAILALFFHMYSHNLSLSKRAQPERMVLVLSWRSSSVGQ	60
QY	61	LFQHPDVFYLMPEPAHVMWTFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGGRQ	120
DB	61	LFQHPDVFYLMPEPAHVMWTFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGGRQ	120
QY	121	SSLFQWENSRLACAPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
DB	121	SSLFQWENSRLACAPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
QY	181	FFNLSQSLYPLKDPISLNHIVLVRDPAVFRSREPTKGLMDISRTVMGHEQKLKED	240
DB	181	FFNLSQSLYPLKDPISLNHIVLVRDPAVFRSREPTKGLMDISRTVMGHEQKLKED	240
QY	241	QPYVVMQVYCQSQLEIKTIGSLPKALQERYLVRYEDLAPAPVAQTSRMTEFYGLEFLP	300
DB	241	QPYVVMQVYCQSQLEIKTIGSLPKALQERYLVRYEDLAPAPVAQTSRMTEFYGLEFLP	300
QY	301	HLQTMVNHITRGKMGDAFHTNARDALNVSAQMRSLPYEKVSRLLKAGGDANLLGYR	360
DB	301	HLQTMVNHITRGKMGDAFHTNARDALNVSAQMRSLPYEKVSRLLKAGGDANLLGYR	360
QY	361	HVRSEQEQRLNLLDLSTWTVEQIH	386
DB	361	HVRSEQEQRLNLLDLSTWTVEQIH	386

RESULT 2

US-09-190-911-1

Sequence 1, Application US/09190911
Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Tangemann, Kirsten
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-10701P
CURRENT APPLICATION NUMBER: US/09/190, 911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045, 284
EARLIER FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-09-190-911-1

Query Match 100.0%; Score 386; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKKMLLFLVSQMAILALFFHMYSHNLSLSKRAQPERMVLVLSWRSSSVGQ	60
DB	1	MLPRKKMLLFLVSQMAILALFFHMYSHNLSLSKRAQPERMVLVLSWRSSSVGQ	60
QY	61	LFQHPDVFYLMPEPAHVMWTFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGGRQ	120
DB	61	LFQHPDVFYLMPEPAHVMWTFKOSTAMLMHAYRDLIRAVFLCDMSVFDAVMPGGRQ	120
QY	121	SSLFQWENSRLACAPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
DB	121	SSLFQWENSRLACAPACDIIPODEIIPRAHCRLLCSQOPEVEVEKACRSYSHVLEKVR	180
QY	181	FFNLSQSLYPLKDPISLNHIVLVRDPAVFRSREPTKGLMDISRTVMGHEQKLKED	240

```
Db 181 FFIQSLYPLIKDPSLNLHIVLVRDPAVFRSRERTKGDLMDSRIVMGOHEOKLKED 240
Qy 241 QPYVMGVICQSOLEITYKITISLPKALOERYLLVREDLARAPVQTSRMYEEVGLFLP 300
Db 241 QPYVMGVICQSOLEITYKITISLPKALOERYLLVREDLARAPVQTSRMYEEVGLFLP 300
Qy 301 HLOTWVHNITRGKMGDHAFTNARDALNYSQAMRWSLPYEKVSRLOKACGDAMNLIQYR 360
Db 301 HLOTWVHNITRGKMGDHAFTNARDALNYSQAMRWSLPYEKVSRLOKACGDAMNLIQYR 360
Qy 361 HVRSSEQRNLDDLSTWTVPEQIH 386
Db 361 HVRSSEQRNLDDLSTWTVPEQIH 386
```

Search completed: February 20, 2003, 13:26:29
Job time : 18 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:22:51 ; Search time 21 Seconds
(without alignments)
1767.042 Million cell updates/sec

Title: US-09-816-825-2
Perfect score: 386
Sequence: 1 MLDPKKMKLLFLVSQMAIL.....EQRNLLDLSTWTPQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result	Query		
No.	Score	Match Length	ID
Description			
No matches found			

Search completed: February 20, 2003, 13:26:05
Job time : 21 secs

This page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:19:55 ; Search time 14 Seconds
(without alignments)
1143.562 Million cell updates/sec

Title: US-09-816-825-2

Perfect score: 386
Sequence: 1 MLDPKMKLLFLVSGMAIL.....EGRNLLDLSTWVPEQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 15

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	ID	Description
---------------	----------------	--------------	----	-------------

No matches found

Search completed: February 20, 2003, 13:24:55
Job time : 14 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:25:41 ; Search time 13 Seconds

(without alignments)
758.605 Million cell updates/sec

Title: US-09-816-825-2

Sequence: 1 MLPRKMKLLFLVSQMAIL.....EQRNLLDLSTWVPEQIH 386

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 140259 seqs, 25548976 residues

Word size: 15

Total number of hits satisfying chosen parameters: 9

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database:

Published Applications_AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
- 2: /cgn2_6/ptodata/1/pubpaa/PCU_NEW_PUB.pep:*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
- 4: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep:*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep:*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	386	100.0	386	9	US-10-007-262-1
2	386	100.0	386	9	US-09-816-825-2
3	60	15.5	171	10	US-09-927-602-8
4	19	4.9	169	10	US-09-927-602-6
5	19	4.9	169	10	US-09-927-602-7
6	19	4.9	390	10	US-09-927-602-4
7	19	4.9	395	10	US-09-927-602-2
8	19	4.9	395	10	US-09-927-602-3
9	19	4.9	418	10	US-09-927-602-5

ALIGNMENTS

RESULT 1
US-10-007-262-1
; Sequence 1, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:

APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/10/007,262
PRIOR FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 386; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MLPRKMKLLFLVSQMAILALFFHMTSHNITSLSKRAQPERMHTVLSWRSGSSFFVGQ	60
DB	1	MLPRKMKLLFLVSQMAILALFFHMTSHNITSLSKRAQPERMHTVLSWRSGSSFFVGQ	60
QY	61	LFQHPDVFYLMPEPAWHVMTFKOSTAMMLHMAVRDLIRAVFLCDMSVFAYMPEGRQ	120
DB	61	LFQHPDVFYLMPEPAWHVMTFKOSTAMMLHMAVRDLIRAVFLCDMSVFAYMPEGRQ	120
QY	121	SSLFQWNSRALCAPACDIIPDEIIPRAHCRLLCSQPEVEVEKACRSYSHVLEKEVR	180
DB	121	SSLFQWNSRALCAPACDIIPDEIIPRAHCRLLCSQPEVEVEKACRSYSHVLEKEVR	180
QY	181	FFNQSLYPLPKDSLNHTYHVLVDRPRAVRSERFKGDLMDSRVYMGQHEQKLEKED	240
DB	181	FFNQSLYPLPKDSLNHTYHVLVDRPRAVRSERFKGDLMDSRVYMGQHEQKLEKED	240
QY	241	QPYVMQVYCQSQLEIKTIQSLPKALQERYLTVREYEDLARAFAQTSRMVEFYGLEFLP	300
DB	241	QPYVMQVYCQSQLEIKTIQSLPKALQERYLTVREYEDLARAFAQTSRMVEFYGLEFLP	300
QY	301	HLQVWEHTITFGKMGDHAFTNARDALNVSQANRWSLPYEKYSRLQACGDANNLLGYR	360
DB	301	HLQVWEHTITFGKMGDHAFTNARDALNVSQANRWSLPYEKYSRLQACGDANNLLGYR	360
QY	361	HVRSEQEQRNLLDLSTWVPEQIH	386
DB	361	HVRSEQEQRNLLDLSTWVPEQIH	386

RESULT 2
US-09-816-825-2
; Sequence 2, Application US/09816825
; Patent No. US20010051370A1
; GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-825-2

```

Query Match          100.0%; Score 386; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 386; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MLPLPKKKLLFLVSMQAILALFFHHYNSNISLSMKAOPEHHVVLSSWSSGSSFFGQ 60
   |||
DB 1 MLPLPKKKLLFLVSMQAILALFFHHYNSNISLSMKAKPERHHVVLSSWSSGSSFFGQ 60

QY 61 LFGQHPDVEFLMEPAHVMWTFKOSTAMMLHMAVRDLIAVFLCDMSVFDAYMEGPRQ 120
   |||
DB 61 LFGQHPDVEFLMEPAHVMWTFKOSTAMMLHMAVRDLIAVFLCDMSVFDAYMEGPRQ 120

QY 121 SSLFOWENSRALCSAPACDIIPDEIIPRAHCHLLCSQGFVEVEACRSYSHVYKEVR 180
   |||
DB 121 SSLFOWENSRALCSAPACDIIPDEIIPRAHCHLLCSQGFVEVEACRSYSHVYKEVR 180

QY 181 FFNLOSLYPLPKDPSLNLIHVLRDPRAVFRSERTKGLMDSRTVNGOHEOKLKKED 240
   |||
DB 181 FFNLOSLYPLPKDPSLNLIHVLRDPRAVFRSERTKGLMDSRTVNGOHEOKLKKED 240

QY 241 QPYVYVQVICOQSOLEIYKTIOS.PKALOERYLLIYVEDIARAPVQTSNMYEFGLEFLP 300
   |||
DB 241 QPYVYVQVICOQSOLEIYKTIOS.PKALOERYLLIYVEDIARAPVQTSNMYEFGLEFLP 300

QY 301 HLOTWNNITRGKMGDHAFTNARDALNVSQAWMSLPEYKYSRLQKACGDAMNLLGYR 360
   |||
DB 301 HLOTWNNITRGKMGDHAFTNARDALNVSQAWMSLPEYKYSRLQKACGDAMNLLGYR 360

QY 361 HVSEOEQRNLLDLSTWVPEQIH 386
   |||
DB 361 HVSEOEQRNLLDLSTWVPEQIH 386

RESULT 3
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8

Query Match          15.5%; Score 60; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 7,3e-50;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 KACRSYSHVYLKEVRFNLOSLYPLKDPSLNLIHVLRDPRAVFRSERTKGLMDS 225
   |||
DB 53 KACRSYSHVYLKEVRFNLOSLYPLKDPSLNLIHVLRDPRAVFRSERTKGLMDS 112

RESULT 4
US-09-927-602-6
; Sequence 6, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien

```

```

; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-6

Query Match          4.9%; Score 19; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGSSFFVGOLF 62
   |||
DB 10 HVLVSSWRSGSSFFVGOLF 28

RESULT 5
US-09-927-602-7
; Sequence 7, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-7

Query Match          4.9%; Score 19; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 1.1e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSYSHVYLKEVRFNLO 185
   |||
DB 51 ACRSYSHVYLKEVRFNLO 69

RESULT 6
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien

```

US-09-927-602-4

Query Match 4.9%; Score 19; DB 10; Length 390;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 167 ACRSSHYVLKEVREFNIO 185

DB 165 ACRSSHYVLKEVREFNIO 183

RESULT 7

US-09-927-602-2
; Sequence 2, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; PRIOR APPLICATION NUMBER: 2001-08-09
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-2

Query Match 4.9%; Score 19; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGGSFVGOLF 62

DB 42 HVLVSSWRSGGSFVGOLF 60

RESULT 8

US-09-927-602-3
; Sequence 3, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; PRIOR APPLICATION NUMBER: 2001-08-09
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: VARIANT
; LOCATION: (1)...(395)
; OTHER INFORMATION: Xaa = any amino acid
US-09-927-602-3

Query Match 4.9%; Score 19; DB 10; Length 395;
Best Local Similarity 100.0%; Pred. No. 2.2e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGGSFVGOLF 62
DB 42 HVLVSSWRSGGSFVGOLF 60

RESULT 9

US-09-927-602-5
; Sequence 5, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; PRIOR APPLICATION NUMBER: 2001-08-09
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-927-602-5

Query Match 4.9%; Score 19; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 2.3e-10;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 44 HVLVSSWRSGGSFVGOLF 62
DB 66 HVLVSSWRSGGSFVGOLF 84

Search completed: February 20, 2003, 13:30:26
Job time : 13 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:11 ; Search time 51.3714 Seconds
(without alignments)
75.222 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166

Sequence: 1 SWRSGSFVQGLFGQHPDYFLMEPRAMHV 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	166	100.0	171 23	AB881560 Human high endothelial cell
2	166	100.0	380 23	AA011274 Human L-selectin s
3	166	100.0	386 23	AAV39918 Human glycosyl sul
4	166	100.0	386 21	AAV92919 Human transferrase
5	166	100.0	388 20	AAW33039 Human polypeptide,
6	166	100.0	388 20	AAV39919 Mouse glycosyl sul
7	160	96.4	169 23	AB881558 Human corneal N-ac
8	160	96.4	395 22	AAV72638 Mouse glycosyl sul
9	160	96.4	395 22	AAV72640 Human glycosyl sul
10	160	96.4	395 23	AB881554 Human corneal N-ac

11	160	96.4	395 23	AB881555 Consensus N-acetyl
12	160	96.4	395 23	AAE15438 Human drug metabol
13	160	96.4	395 23	AAU11275 Murine intestinal-
14	160	96.4	418 23	AB881557 Mouse intestinal N
15	157	94.6	169 23	AB881559 Human intestinal N
16	157	94.6	390 22	AAV72639 Human glycosyl sul
17	157	94.6	390 22	AB881556 Human glycosyl sul
18	157	94.6	418 21	AB881561 Human intestinal N
19	157	94.6	418 21	AB881562 Human ORF ORP1711
20	157	94.6	418 21	AB881563 Human N-acetylgluc
21	157	94.6	418 21	AB881564 Mouse N-acetylgluc
22	157	94.6	418 21	AB881565 Human N-acetylgluc
23	157	94.6	418 21	AB881566 Human N-acetylgluc
24	157	94.6	418 21	AB881567 Human N-acetylgluc
25	157	94.6	418 21	AB881568 Human N-acetylgluc
26	157	94.6	418 21	AB881569 Human N-acetylgluc
27	157	94.6	418 21	AB881570 Human N-acetylgluc
28	157	94.6	418 21	AB881571 Human N-acetylgluc
29	157	94.6	418 21	AB881572 Human N-acetylgluc
30	157	94.6	418 21	AB881573 Human N-acetylgluc
31	157	94.6	418 21	AB881574 Human N-acetylgluc
32	157	94.6	418 21	AB881575 Human N-acetylgluc
33	157	94.6	418 21	AB881576 Human N-acetylgluc
34	157	94.6	418 21	AB881577 Human N-acetylgluc
35	157	94.6	418 21	AB881578 Human N-acetylgluc
36	157	94.6	418 21	AB881579 Human N-acetylgluc
37	157	94.6	418 21	AB881580 Human N-acetylgluc
38	157	94.6	418 21	AB881581 Human N-acetylgluc
39	157	94.6	418 21	AB881582 Human N-acetylgluc
40	157	94.6	418 21	AB881583 Human N-acetylgluc
41	157	94.6	418 21	AB881584 Human N-acetylgluc
42	157	94.6	418 21	AB881585 Human N-acetylgluc
43	157	94.6	418 21	AB881586 Human N-acetylgluc
44	157	94.6	418 21	AB881587 Human N-acetylgluc
45	157	94.6	418 21	AB881588 Human N-acetylgluc
46	157	94.6	418 21	AB881589 Human N-acetylgluc
47	157	94.6	418 21	AB881590 Human N-acetylgluc
48	157	94.6	418 21	AB881591 Human N-acetylgluc
49	157	94.6	418 21	AB881592 Human N-acetylgluc
50	157	94.6	418 21	AB881593 Human N-acetylgluc
51	157	94.6	418 21	AB881594 Human N-acetylgluc
52	157	94.6	418 21	AB881595 Human N-acetylgluc
53	157	94.6	418 21	AB881596 Human N-acetylgluc
54	157	94.6	418 21	AB881597 Human N-acetylgluc
55	157	94.6	418 21	AB881598 Human N-acetylgluc
56	157	94.6	418 21	AB881599 Human N-acetylgluc
57	157	94.6	418 21	AB881600 Human N-acetylgluc
58	157	94.6	418 21	AB881601 Human N-acetylgluc
59	157	94.6	418 21	AB881602 Human N-acetylgluc
60	157	94.6	418 21	AB881603 Human N-acetylgluc
61	157	94.6	418 21	AB881604 Human N-acetylgluc
62	157	94.6	418 21	AB881605 Human N-acetylgluc
63	157	94.6	418 21	AB881606 Human N-acetylgluc
64	157	94.6	418 21	AB881607 Human N-acetylgluc
65	157	94.6	418 21	AB881608 Human N-acetylgluc
66	157	94.6	418 21	AB881609 Human N-acetylgluc
67	157	94.6	418 21	AB881610 Human N-acetylgluc
68	157	94.6	418 21	AB881611 Human N-acetylgluc
69	157	94.6	418 21	AB881612 Human N-acetylgluc
70	157	94.6	418 21	AB881613 Human N-acetylgluc
71	157	94.6	418 21	AB881614 Human N-acetylgluc
72	157	94.6	418 21	AB881615 Human N-acetylgluc
73	157	94.6	418 21	AB881616 Human N-acetylgluc
74	157	94.6	418 21	AB881617 Human N-acetylgluc
75	157	94.6	418 21	AB881618 Human N-acetylgluc
76	157	94.6	418 21	AB881619 Human N-acetylgluc
77	157	94.6	418 21	AB881620 Human N-acetylgluc
78	157	94.6	418 21	AB881621 Human N-acetylgluc
79	157	94.6	418 21	AB881622 Human N-acetylgluc
80	157	94.6	418 21	AB881623 Human N-acetylgluc
81	157	94.6	418 21	AB881624 Human N-acetylgluc
82	157	94.6	418 21	AB881625 Human N-acetylgluc
83	157	94.6	418 21	AB881626 Human N-acetylgluc

230	42.5	25.6	210	23	ABP28214	303	42	25.3	526	22	ABB60414	Drosophila melanog
231	42.5	25.6	229	21	AAE67174	529	42	25.3	23	AAU74387	Breast tumour-spec	
232	42.5	25.6	229	21	ABE52377	304	42	25.3	23	AAU74387	Breast tumour-spec	
233	42.5	25.6	304	23	ABAB4806	305	42	25.3	531	20	AAU25979	Mistletoe lectin I
234	42.5	25.6	348	22	AAAB3467	306	42	25.3	532	20	AAU25982	Mistletoe lectin I
235	42.5	25.6	453	22	AAAB8601	307	42	25.3	545	22	AAU40860	Human polypeptide
236	42.5	25.6	514	21	AAAG54762	308	42	25.3	544	18	AAU10021	Prepro mistletoe I
237	42.5	25.6	514	21	ABP29967	309	42	25.3	564	20	AAU90127	Mistletoe lectin p
238	42.5	25.6	604	23	ABP25463	310	42	25.3	616	20	AAU42287	Human anchoreless p
239	42.5	25.6	614	23	AAU67485	311	42	25.3	616	20	AAU27399	Human P450 reducta
240	42.5	25.6	657	22	ABAB6846	312	42	25.3	621	21	AAU19285	Human P450 reducta
241	42.5	25.6	823	21	AAU75441	313	42	25.3	626	21	AAU28629	A polypeptide with
242	42.5	25.6	871	21	AAAG49343	314	42	25.3	656	21	AAU82018	Human B1Agl1 antiq
243	42.5	25.6	875	21	AAAG49322	315	42	25.3	656	22	AAU65977	Human immunogenic
244	42.5	25.6	877	21	AAAG49321	316	42	25.3	656	22	AAU65977	B305D isoform C sp
245	42.5	25.6	960	23	ABAB3642	317	42	25.3	656	22	AAU65977	Human B1Agl1 antiq
246	42.5	25.6	2304	22	ABAB60219	318	42	25.3	656	22	AAU65977	Human prostate-spe
247	42.5	25.6	2304	22	ABAB60219	319	42	25.3	656	22	AAU65977	Human prostate-spe
248	42	25.3	61	21	AAAG22574	320	42	25.3	656	22	AAU74816	Human prostate-spe
249	42	25.3	89	22	AAU13619	321	42	25.3	656	23	AAU74816	Prostate tumour an
250	42	25.3	95	23	ABAB80785	322	42	25.3	671	21	AAU74378	Breast tumour-spec
251	42	25.3	99	22	AAAB4777	323	42	25.3	671	21	AAU82019	Human B1Agl1 antiq
252	42	25.3	108	21	AAU02246	324	42	25.3	671	21	AAU82019	Human immunogenic
253	42	25.3	111	22	AAU04889	325	42	25.3	671	22	AAU65978	B305D isoform C sp
254	42	25.3	135	22	AAU59289	326	42	25.3	671	22	AAU65979	Human prostate-spe
255	42	25.3	151	22	ABG21145	327	42	25.3	671	22	AAU01134	Human prostate-cdn
256	42	25.3	163	23	ABAB60948	328	42	25.3	671	22	AAU99019	Human prostate-spe
257	42	25.3	170	20	AAAB89065	329	42	25.3	671	22	AAU74817	Human prostate-spe
258	42	25.3	170	20	ABAB51236	330	42	25.3	671	23	AAU74379	Human B305D splice
259	42	25.3	217	23	ABAB60947	331	42	25.3	676	20	AAU74286	Breast tumour-spec
260	42	25.3	225	22	AAAB96159	332	42	25.3	676	20	AAU72398	Human P450 reducta
261	42	25.3	250	20	AAAG4284	333	42	25.3	676	23	AAU17538	Human P450 reducta
262	42	25.3	251	23	ABAB09920	334	42	25.3	876	22	AAU17538	Human P450 reducta
263	42	25.3	252	19	AAAB4659	335	42	25.3	1019	23	AAU37030	Human P450 reducta
264	42	25.3	252	19	AAAB4659	336	42	25.3	1354	22	ABAB92447	Staphylococcus aur
265	42	25.3	253	18	AAU10022	337	42	25.3	1719	21	AAU82017	Herbicidally activ
266	42	25.3	253	20	AAAB90135	338	42	25.3	1719	22	AAU69777	Drosophila melanog
267	42	25.3	254	20	AAU25980	339	42	25.3	1719	22	AAU01132	Human immunogenic
268	42	25.3	254	20	AAU25983	340	42	25.3	1719	22	AAU01132	Human prostate-cdn
269	42	25.3	258	22	ABAB68035	341	42	25.3	1719	22	AAU99017	Human prostate-spe
270	42	25.3	260	22	AAAB1774	342	42	25.3	1719	22	AAU74815	Human prostate-spe
271	42	25.3	268	22	AAU78652	343	42	25.3	1719	22	AAU74815	Human prostate-spe
272	42	25.3	268	22	AAU39537	344	41.5	25.0	53	23	AAAB33798	Human B305D splice
273	42	25.3	273	23	ABAB97278	345	41.5	25.0	56	22	AAAB07562	Protein encoded by
274	42	25.3	273	23	AAAG26632	346	41.5	25.0	90	22	AAU2597	Human haematologic
275	42	25.3	328	21	AAAG26632	347	41.5	25.0	90	22	AAU2597	Novel human colon
276	42	25.3	328	22	AAU79636	348	41.5	25.0	105	22	AAU92571	Human digestive sy
277	42	25.3	328	22	AAU41333	349	41.5	25.0	105	22	ABAB27682	Human peptide #333
278	42	25.3	335	21	AAAB42940	350	41.5	25.0	105	22	ABAB32852	Peptide #338 encod
279	42	25.3	335	21	AAAB42943	351	41.5	25.0	105	22	AAAB18334	Human brain expres
280	42	25.3	335	21	AAU939376	352	41.5	25.0	105	22	AAU53655	Human bone marrow
281	42	25.3	335	22	AAU94899	353	41.5	25.0	105	22	AAU66038	Peptide #311 encod
282	42	25.3	335	22	AAAB8486	354	41.5	25.0	105	22	AAU13907	Peptide #332 encod
283	42	25.3	335	22	AAAB66125	355	41.5	25.0	105	22	AAU01650	Human polypeptide
284	42	25.3	335	23	ABAB5517	356	41.5	25.0	105	22	ABAB3586	Human polypeptide
285	42	25.3	335	23	ABAB4911	357	41.5	25.0	105	22	AAU07803	Staphylococcus epi
286	42	25.3	335	23	AAU83676	358	41.5	25.0	182	17	AAAB37997	Blowfly larvae PM9
287	42	25.3	378	21	AAAB3212	359	41.5	25.0	198	17	AAAB85214	Human immunohaema
288	42	25.3	384	22	AAAB8628	360	41.5	25.0	246	22	AAAB9436	Putative p. abysal
289	42	25.3	384	22	AAAG5976	361	41.5	25.0	334	22	AAAB74771	Human polypeptide
290	42	25.3	384	23	ABAB0999	362	41.5	25.0	407	22	AAAB40808	Human polypeptide
291	42	25.3	384	23	ABAB1001	363	41.5	25.0	437	20	AAAB88342	Salmonella enteric
292	42	25.3	384	23	AAU74377	364	41.5	25.0	446	22	AAU39022	Human polypeptide
293	42	25.3	391	22	ABAB16370	365	41.5	25.0	446	22	AAU62275	Human gene 17-enco
294	42	25.3	435	20	AAU42288	366	41.5	25.0	455	22	AAU03202	Human gene 1-enco
295	42	25.3	435	20	AAU42288	367	41.5	25.0	455	22	ABAB64398	Human albumin fusi
296	42	25.3	490	22	ABAB27400	368	41.5	25.0	458	22	AAU74771	Human secreted pro
297	42	25.3	492	22	ABAB6289	369	41.5	25.0	480	17	AAAB85213	Blowfly larvae PM9
298	42	25.3	492	22	AAAB1914	370	41.5	25.0	508	22	AAU664726	Murine transcripti
299	42	25.3	501	22	AAU43517	371	41.5	25.0	553	22	ABAB60817	Drosophila melanog
300	42	25.3	501	22	AAU43620	372	41.5	25.0	582	22	AAU62276	Human gene 17-enco
301	42	25.3	501	22	AAU73855	373	41.5	25.0	597	21	AAU30765	Arabidopsis thalia
302	25.3	25.3	517	22	AAU65240	374	41.5	25.0	645	21	AAU45573	Human albumin fusi
						375	41.5	25.0	667	23	ABAB63572	Human albumin fusi

960	39	23.5	528	23	AAU72955	Neisseria meningit
961	39	23.5	536	22	ABE62486	Drosophila melanog
962	39	23.5	544	22	AAE09716	Human extracellular
963	39	23.5	544	23	ABP27253	Streptococcus poly
964	39	23.5	544	23	AAU72718	Human kinase h1418
965	39	23.5	547	23	AAE14507	Human carboxyleste
966	39	23.5	557	23	AAU78802	Human protein S60
967	39	23.5	558	22	AAU03526	Human protein kina
968	39	23.5	568	22	ABG21986	Novel human diago
969	39	23.5	571	21	AAI18914	A novel polypeptid
970	39	23.5	571	21	AAU71107	Human hydrolyase pr
971	39	23.5	571	22	AAU12442	Human PRO1887 poly
972	39	23.5	571	22	AAE04101	Human gene 2 encod
973	39	23.5	571	23	ABE64341	Human albumin fusi
974	39	23.5	571	23	ABE84949	Human PRO1887 prot
975	39	23.5	572	23	AAU83696	Human PRO protein,
976	39	23.5	572	23	ABH77576	Physcomitrella pat
977	39	23.5	572	23	ABE55555	Human angiotensin
978	39	23.5	574	22	ABE58877	Drosophila melanog
979	39	23.5	580	19	AAU11093	H. pylori ORF hp6p
980	39	23.5	581	22	AAU33503	Novel human secret
981	39	23.5	583	19	AAU71506	Helicobacter polyp
982	39	23.5	593	22	AAE46350	H. pylori HPN191 p
983	39	23.5	605	21	AAE26942	Soybean auxin tran
984	39	23.5	606	22	ABE52567	Escherichia coli p
985	39	23.5	610	21	AAE53795	Archidopsis thalia
986	39	23.5	612	22	AAU79786	Human protein S60
987	39	23.5	621	22	AAE94433	Human protein sequ
988	39	23.5	708	22	AAU48037	Proionibacterium
989	39	23.5	711	19	AAW55103	Streptococcus pneu
990	39	23.5	711	23	ABP54597	S. pneumoniae SP05
991	39	23.5	717	22	ABE61453	Drosophila melanog
992	39	23.5	721	22	ABE15463	Novel human diago
993	39	23.5	727	20	AAU27127	Human brain tissue
994	39	23.5	777	21	AAU99427	Human PRO1491 (ONQ
995	39	23.5	777	22	AAU29197	Human PRO polypept
996	39	23.5	777	22	AAE66176	Protein of the inv
997	39	23.5	783	19	AAW64560	Candida albicans o
998	39	23.5	783	21	AAE96164	Candida albicans o
999	39	23.5	800	22	ABE16023	Novel human diago
1000	39	23.5	807	22	AAE62729	Amino acid sequenc

ALIGNMENTS

RESULT 1
ID ABB81560 standard; Protein; 171 AA.
AC ABB81560;
DT 05-SEP-2002 (first entry)
XX
DE Human high endothelial cell G1CNAc6ST portion SMO ID NO:8.
XX
KW Human; N-acetylglucosamine-6-sulfotransferase; enzyme; G1CNAc6ST;
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological.
XX
OS Homo sapiens.
XX
PN US2002061562-A1.
XX
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU/) FUKUDA M N.
PA (AKAM/) AKAMA T O.
XX

PI Fukuda MN, Akama TO;
XX WPI: 2002-507643/54.
DR
XX
PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy
XX
PS Example 2; Fig 3C; 69pp; English.
XX
CC The present invention describes human corneal
CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
CC and has ophthalmological activity. (I) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (I) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents a portion of the human high endothelial cell
CC N-acetylglucosamine-6-sulfotransferase (G1CNAc6ST), which is given in
CC the exemplification of the present invention.
XX
SQ Sequence 171 AA:

Query Match 100.0%; Score 166; DB 23; Length 171;
Best Local Similarity 100.0%; Pred. No. 7.6e-18;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSSSEFVGQLFQHPDVFYLMPEAWHY 29
DB 19 SWRSSSEFVGQLFQHPDVFYLMPEAWHY 47

RESULT 2
ID AAU11274 standard; Protein; 380 AA.
AC AAU11274;
DT 12-MAR-2002 (first entry)
XX
DE Human L-selectin sulfotransferase-2 (LSTF-2) protein.
XX
KW Human; beta1,3GNT; beta1,3-N-acetylglucosaminyltransferase; MECA-79;
KW L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;
KW ulcerative colitis; inflammatory skin disorder; psoriasis; lichen planus;
KW allergic contact dermatitis; lymphoma; chronic pneumonia; LSTF-2;
KW delayed-type hypersensitivity reaction; hyperplastic thymus; antileuk;
KW antiinflammatory; antipsoriatic; antidiabetic; dermatological;
KW antiallergic.
XX
OS Homo sapiens.
XX
PN WO200185177-A1.
XX
PD 15-NOV-2001.
XX
PF 10-MAY-2001; 2001WO-US15452.
XX
PR 11-MAY-2000; 2000US-0569320.
XX
PA (BORN-) BORNHAM INST.
XX
PA Fukuda M, Yeh J, Hiraoka N;
XX
DR WPI: 2002-075226/10.
XX
DR N-PSDB; AAS16947.
XX

New enzyme, useful for modifying acceptor molecule, comprises an
PT isolated L-selectin sulfotransferase-2 that directs expression of
PT L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or
PT intestinal G1CNAc 6-sulfotransferase

PS Claim 21; Fig 4; 98pp; English.

XX
CC The present invention provides a method of modifying an acceptor molecule
CC by contacting the acceptor with an isolated
CC beta1,3-N-acetylglucosaminyltransferase (beta1,3GNT) or an active
CC fragment, where beta1,3GNT directs expression of a MECA-79 antigen. The
CC invention also provides a method of treating or preventing an
CC L-selectin-mediated condition by reducing the expression or activity of a
CC beta1,3GNT that directs expression of a MECA-79 antigen. This can be done
CC by administering to the subject an oligosaccharide L-selectin antagonist
CC that inhibits binding of L-selectin to a MECA-79 antigen, for example by
CC administering antibody material that specifically binds beta1,3GNT,
CC and/or a beta1,3GNT antisense nucleic acid molecule. L-selectin
CC sulfotransferase-2 (LST-2) also directs MECA-79 antigen expression.
CC Alternatively, the expression or activity of LST-2 or its active
CC fragment can be reduced in combination with reducing the expression or
CC activity of beta1,3GNT. The method is useful for treating L-selectin
CC mediated conditions such as Crohn's disease and ulcerative colitis,
CC inflammatory disorders of the skin such as allergic contact dermatitis,
CC psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type
CC hypersensitivity reactions, diabetes and hyperplastic thymus. This
CC sequence represents human LST-2.

XX
SQ Sequence 380 AA;

Query Match 100.0%; Score 166; DB 23; Length 380;
Best Local Similarity 100.0%; Pred. No. 1,9e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SWRSGSFFVGQLFGQHPDVFYLMPPAMHV 29
|||||
DB 44 SWRSGSFFVGQLFGQHPDVFYLMPPAMHV 72
|||||

RESULT 3
AAV39918
ID AAV39918 standard; Protein: 386 AA.
XX
AC AAV39918;
XX
DT 08-DEC-1999 (first entry)
XX
DE Human glycosyl sulfotransferase-3 protein sequence.
XX
KW Glycosyl sulfotransferase; GST-3; detection: diagnosis; leukocyte homing;
KW selectin binding interaction; inflammation; lymphocyte homing; human;
KW secondary lymph organ.
XX
OS Homo sapiens.
XX
PN WO9949018-A1.
XX
PD 30-SEP-1999.
XX
PE 26-FEB-1999; 99WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
XX
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIV CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
XX
DR WPT: 1999-580442/49.
DR N-PsDB; AA220792.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
PT polynucleotides
XX
PS Claim 2; Fig 1; 59pp; English.
XX
CC This sequence is the human glycosyl sulfotransferase-3 (GST-3) of

CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC These sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions,
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.

XX
SQ Sequence 386 AA;

Query Match 100.0%; Score 166; DB 20; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SWRSGSFFVGQLFGQHPDVFYLMPPAMHV 29
|||||
DB 50 SWRSGSFFVGQLFGQHPDVFYLMPPAMHV 78
|||||

RESULT 4
AAV79219
ID AAV79219 standard; Protein: 386 AA.
XX
AC AAV79219;
XX
DT 19-JUN-2000 (first entry)
XX
DE Human transferase TRNSFS-11.
XX
KW Transferase; TRNSFS-11; human; antitumor; cell proliferation;
KW gastrointestinal disorder; developmental disorder;
KW genetic disorder; neurological disorder; reproductive disorder;
KW smooth muscle disorder; immunological disorder; inflammation;
KW diagnosis; therapy; N-acetylglucosamine 6-O-sulfotransferase.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FH Modified-site 121
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 107
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 217
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 252
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 364
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 380
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 35
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 50
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 81
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 287
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 243
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 30
FT Modified-site /note= "potential O-phosphorylation"
FT Modified-site 308
FT Modified-site /note= "potential N-glycosylation"
FT Modified-site 329
FT Modified-site /note= "potential N-glycosylation"
FT Domain 7..23
FT Modified-site /note= "transmembrane domain"
XX
PN WO200014251-A2.
XX
PD 16-MAR-2000.

PF 09-SEP-1999; 99WO-US20989.
XX
PR 10-SEP-1998; 98US-0150657.
PR 04-NOV-1998; 98US-0186779.
PR 11-MAY-1999; 99US-0133642.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
PI Hillman JL, Azimzai Y;
XX
DR WPI, 2000-25696/22.
DR N-PSDB; AA294211.
XX
PT Human transferase proteins useful for preventing, diagnosing and
PT treating cancers and developmental, gastrointestinal, genetic,
PT immunological, neurological, reproductive and smooth muscle disorders -
XX
PS Claim 1; Page 90-91; 113pp; English.
XX
CC The present sequence is that of human transferase TRNSFS-11, 1 of
CC 15 claimed novel human transferase proteins of the invention (see
CC AAY79209-23). The sequence was deduced from a cDNA clone (see
CC AA294211) isolated from a gallbladder library. It shows homology to
CC mouse N-acetylglucosamine 6-O-sulfotransferase. TRNSFS-11 is
CC expressed in dermatologic and gastrointestinal tissues,
CC especially those associated with inflammation and cell
CC proliferation. The new human transferases and polynucleotides can
CC be used in the diagnosis, prevention and treatment of cancer,
CC developmental disorders, gastrointestinal disorders, genetic
CC disorders, immunological disorders, neurological disorders,
CC reproductive disorders, and smooth muscle disorders. The
CC polypeptides can also be used to raise antibodies, and to screen
CC for agonists and antagonists of transferase activity.
XX
SQ Sequence 386 AA;

Query Match 100.0%; Score 166; DB 21; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVGLFGQHPDYFLMEPAMHV 29
Db 50 SMRSGSFVGLFGQHPDYFLMEPAMHV 78
|||||
RESULT 5
AA93309
ID AAM93309 standard; Protein; 386 AA.
XX
AC AAM93309;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide, SEQ ID NO: 2817.
XX
KM Human; full length cDNA; cDNA synthesis; oligo-capping.
XX
OS Homo sapiens.
XX
PN EP130094-A2.
XX
PD 05-SEP-2001.
XX
PF 07-JUL-2000; 2000EP-0114089.
XX
PR 08-JUL-1999; 99JP-0194486.
PR 11-JAN-2000; 2000JP-0118774.
PR 02-MAY-2000; 2000JP-0183765.
XX
PA (HELT-) HELIX RES INST.
XX
PI Ota T, Nishiyawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
XX
DR WPI, 2001-524255/58.
DR N-PSDB; AA94429.
XX
PT 830 Primers useful for synthesizing full length cDNA clones and their
PT use in genetic manipulation -
XX
PS Claim 8; SEQ ID NO 2817; 1380pp + sequence listing; English.
XX
CC The invention relates to primers for synthesizing full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5' and 3' ends of the cDNA
CC molecules have been determined. Primers for synthesizing the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.
XX
SQ Sequence 386 AA;

Query Match 100.0%; Score 166; DB 22; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVGLFGQHPDYFLMEPAMHV 29
Db 50 SMRSGSFVGLFGQHPDYFLMEPAMHV 78
|||||
RESULT 6
AA93919
ID AAY39919 standard; Protein; 388 AA.
XX
AC AAY39919;
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 protein sequence.
XX
KM Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KM selectin binding interaction; inflammation; lymphocyte homing; mouse;
KM secondary lymph organ.
XX
OS Mus sp.
XX
PN W09949018-A1.
XX
PD 30-SEP-1999.
XX
PF 26-FEB-1999; 99WO-US04316.
XX
PR 20-MAR-1998; 98US-0045284.
PR 12-NOV-1998; 98US-0190911.
XX
PA (REGC) UNIV CALIFORNIA.
PA (SYNT) SYNTX USA INC.
XX
PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
XX
DR WPI, 1999-580442/49.
DR N-PSDB; AA20793.
XX
PT Human and murine glycosyl sulfotransferase 3 and related
PT polynucleotides -
XX
PS Claim 2; Fig 4; 59pp; English.
XX
CC This sequence is the mouse glycosyl sulfotransferase-3 (GST-3) of

CC the invention. The nucleic acid sequences, probes and primers derived
CC from these, proteins and antibodies are useful in detecting homologues.
CC The sequences, antibodies and methods are useful in the diagnosis and
CC treatment of diseases associated with selectin binding interactions,
CC including conditions associated with or resulting from the homing of
CC leukocytes to sites of inflammation and the normal homing of lymphocytes
CC to secondary lymph organs.
CC
XX

Sequence 388 AA;

Query Match

Best Local Similarity 100.0%; Score 166; DB 20; Length 388;

Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 29
DB 49 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 77

RESULT 7

ABB81558

ID ABB81558 standard; Protein: 169 AA.

AC ABB81558;

XX

DT 05-SEP-2002 (first entry)

XX

DE Human corneal N-acetylglucosamine-6-sulfotransferase portion SEQ.6.

XX

XX Human: N-acetylglucosamine-6-sulfotransferase; enzyme: GLCNA6ST;

KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;

KM opthalmological.

XX

OS Homo sapiens.

XX

PN US2002061562-A1.

XX

PD 23-MAY-2002.

XX

PF 09-AUG-2001; 2001US-0927602.

XX

PR 11-AUG-2000; 2000US-325773P.

XX

PA (FUKU/) FUKUDA M N.

XX

PI (AKAM/) AKAMA T O.

XX

DR Fukuda MN, Akama TO;

XX

WPI; 2002-507643/54.

PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy -
XX
XX Example 2; Fig 3C; 69pp; English.

CC The present invention describes human corneal
CC N-acetylglucosamine-6-sulfotransferase (1), which is able to catalyse
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (1) is located to chromosome 16q22,
CC and has ophthalmological activity. (1) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (1) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents a portion of the human corneal N-acetylglucosamine-
CC 6-sulfotransferase (GLCNA6ST), which is given in the exemplification
CC of the present invention.
CC
XX

Sequence 169 AA;

Query Match

Best Local Similarity 96.4%; Score 160; DB 23; Length 169;

Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 29
DB 16 SWRSGSFFVQGLFGQHPDVFYLMPEAMHV 44

RESULT 8

AAY72638

ID AAY72638 standard; Protein: 395 AA.

XX

AC AAY72638;

XX

DT 02-MAY-2001 (first entry)

XX

DE Mouse glycosyl sulfotransferase-4 (GST-4).

XX

XX Mouse; glycosyl sulfotransferase-4; GST-4; immunosuppressive;

KW therapy; selectin binding inhibitor; gene therapy; inflammation;

KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;

KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;

KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;

KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;

KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;

KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;

KW asthma; hypersensitivity; rheumatic fever; tissue rejection;

XX

XX Mus musculus.

XX

PN W0200106015-A1.

XX

PD 25-JAN-2001.

XX

PF 19-JUL-2000; 2000MO-US19741.

XX

PR 20-JUL-1999; 99US-0144694.

XX

PR 13-JUL-2000; 2000US-0593828.

XX

PA (REGC) UNIV CALIFORNIA.

XX

PI Rosen SD, Lee JK, Hemmerich S;

XX

PT New glycosyl sulfotransferases (GST)-alpha, GST-beta and GST-6 for
PT diagnostic and therapeutic agent screening applications -
XX
XX Claim 3; Fig 2; 128pp; English.

CC The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4
CC gene is found on chromosome 8E1.
CC GST is a type 2 membrane protein useful for inhibiting a binding event
CC between a selectin and a selectin ligand, which comprises contacting the
CC selectin with a non-sulphated selectin ligand, GST and a small molecular
CC agent that inhibits the sulphation activity of GST. GST is also useful
CC in inhibiting a selectin mediated binding event. GST is useful in gene
CC therapy to treat disorders such as acute or chronic inflammation,
CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
CC disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
CC during transplantation.
CC
XX

Sequence 395 AA;

Query Match

Best Local Similarity 96.4%; Score 160; DB 22; Length 395;

Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSSVVGQLFGQHPDVEYLMPEAMHV 23
48 SWRSGSSVVGQLFGQHPDVEYLMPEAMHV 76

RESULT 9
AAV72640

AC AAY72640;

DT 02-MAY-2001 (first entry)

Human glucosyl sulfotransferase-4beta (GST-4beta)

KM Human: glycosyl sulfotransferase-4beta; GST-4beta; immunosuppressive;
 KM therapy; selectin binding inhibitor; gene therapy; inflammation;
 KM systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 KM polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 KM glomerulonephritis; masthenia gravis; Sjogren's syndrome; adenitis;
 KM Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
 KM demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 KM myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 KM hypersensitivity; rheumatic fever; tissue rejection;
 KM chromosome 16q23.1.

OS Homo sapiens.

PN W0200106015-A1.

PD 25-JAN-2001.

PF 19-JUL-2000; 2000WO-US19741

PR 20-JUL-1999; 99US-0144694.

XX

XX

XX

DR N-PSDB; AAD02697, AAD02700.

PT New glycosyl sulfotransferase

XX
PS Claim 3; Fig 4B; 128pp; English.

CC The present sequence is human glycosyl sulfotransferase-4beta (GST-4
CC beta). GST-4 gene is found on chromosome 16q33.1.
CC GST is a type 2 membrane protein useful for inhibiting a binding event
CC between a selectin and a selectin ligand, which comprises contacting the
CC selectin with a non-sulphated selectin ligand, GST and a small molecular
CC agent that inhibits the sulphonation activity of GST. GST is also useful
CC in inhibiting a selectin mediated binding event. GST is useful in gene
CC therapy to treat disorders such as acute or chronic inflammation,
CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarthritis
CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
CC disease, Grave's disease, adrenailitis, hypoparathyroidism, peritonitis
CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis
CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
CC during transplantation.

Sequence 395 AA.

Query Match	96.48;	Score 160;	DB 22;	Length 395;
Best Local Similarity	96.68;	Pred. No. 1.8e-16;		
Matches 28; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	1	SWSRSSSFVGQLFGQHDPDVFYLMEPAWHV	29
Db	48	SWSRGSSEFVGQLFNQHPDVFYLMEPAWHV	76

RESULT 10
ABB81554

AC ABB81554;
VV

DT 05-SEP-2002 (first entry)

DE Human corneal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:2.

KW Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological; chromosome 16q22.

OS Homo sapiens.

PN US2002061562-A1.

PD 23-MAY-2002

PF 09-AUG-2001; 2001US-0927602

PR 11-AUG-2000; 2000US-325773P

PA (FUKU/) FUKUDA M N

XX

XX

DR N-PSDB; ABN89506.

PT New nucleic acid endo

PI New nucleic acid encoding conaseal
 PI N-acetylglucosamine-6-sulfotransferase, useful for treatment
 PI monitoring and diagnosis of macular corneal dystrophy -
 XX
 PS Claim 13; Fig 1A-D; 69pp; English.

CC The present sequence represents human corneal
CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (I) is located to chromosome 10q22,
CC and has optical susceptibility activity. (I) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (I) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy.

5Q Sequence 395 AA;

Query Match	96.4%	Score 160;	DB 23;	length 395;
Best Local Similarity	96.6%	Pred No. 1.8e-16;		
Matches 28; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

```
QY      1  SWRSGSSFVGQLFGQHPDVFYLMEPAMHY 23
         |||||
Db      48  SWRSGSSFVGQLFQHPDVFYLMEPAMHY 76
```

RESULT 11
ABB81555
ID ABB81555 standard; Protein; 395 AA

AC ABB81555

DT 05-SEP-2002 (first entry

DE Consensus N-acetylglucosamine-6-sulfotransferase SEQ ID NO:3

XX Human; N-acetylglucosamine-6-sulfotransferase; enzyme; G1CNAC6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 OS Mus musculus.
 OS Synthetic.
 XX
 FH Key location/Qualifiers
 FT Misc-difference 10
 FT /label= Ala, Thr, Val
 FT Misc-difference 13
 FT /label= Ala, Val, Ser
 FT Misc-difference 20
 FT /label= Phe, Cys, Gly
 FT Misc-difference 39
 FT /label= Ala, Asp, Glu
 FT Misc-difference 96
 FT /label= Val, Met, Ile
 FT Misc-difference 142
 FT /label= Ala, Thr, Asn
 FT Misc-difference 147
 FT /label= Ala, Asp, Glu
 FT Misc-difference 159
 FT /label= Thr, Ser, Gly
 FT Misc-difference 238
 FT /label= Gly, His, Arg
 FT Misc-difference 294
 FT /label= Ser, Thr, Gly
 FT Misc-difference 371
 FT /label= Ala, Thr, Ser
 FT Misc-difference 380
 FT /label= Leu, Pro, Met
 FT Misc-difference 382
 FT /label= Gly, His, Ser
 FT Misc-difference 384
 FT /label= Thr, Ser, Lys
 FT Misc-difference 390
 FT /label= Ala, Glu
 FT Misc-difference 391
 FT /label= Ser, Lys
 FT Misc-difference 392
 FT /label= His, Gln
 FT Misc-difference 394
 FT /label= Arg, Glu
 FT Misc-difference 395
 FT /label= Asn, Ser
 FT
 XX US2002061562-A1.
 XX
 XX 23-MAY-2002.
 XX
 XX 09-AUG-2001; 2001US-0927602.
 XX
 XX 11-AUG-2000; 2000US-325773P.
 XX
 XX (FUKU/) FUKUDA M N.
 XX (AKAM/) AKAMA T O.
 XX
 XX Fukuda MN, Akama TO;
 XX
 XX WPI; 2002-507643/54.
 XX
 XX New nucleic acid encoding corneal
 XX N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 XX monitoring and diagnosis of macular corneal dystrophy -
 XX Example 5; Fig 2A-B; 69pp; English.
 XX
 XX The present invention describes human corneal
 XX N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 XX sulfation of keratan sulfate (KS). Also described is a method for

CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a consensus N-acetylglucosamine-6-sulfotransferase
 CC which is given in the exemplification of the present invention.
 CC
 XX
 SO Sequence 395 AA;
 QY
 Db 48 SMRSGSFVGOITFGCHPDVFTLMEPMWV 76
 ||||||||||||||||||||||||||||
 1 SMRSGSFVGOITFGCHPDVFTLMEPMWV 29
 ||||||||||||||||||||||||||||
 48 SMRSGSFVGOITFGCHPDVFTLMEPMWV 76
 RESULT 12
 ID AAE15438 standard; Protein; 395 AA.
 XX
 AC AAE15438;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Human drug metabolizing enzyme (DME)-5.
 XX
 KW Human; drug metabolizing enzyme; gene therapy; autoimmune disorder;
 KW inflammatory disorder; acquired immune deficiency syndrome; infection;
 KW AIDS; adult respiratory distress syndrome; cell proliferative disorder;
 KW allergy; anaemia; conjunctivitis; actinic keratosis; arteriosclerosis;
 KW cancer; endocrine disorder; hypochalimus disorder; pituitary disorder;
 KW gastrointestinal disorder; metabolic disorder; developmental disorder;
 KW liver disorder; iritis; cystic fibrosis; Addison's disease; retinitis;
 KW goitre; diabetes; dysphagia; gastric carcinoma; anorexia; eye disorder;
 KW DME-5.
 XX
 OS Homo sapiens.
 XX
 FH Key location/Qualifiers
 FT Peptide 1..32
 FT /label= Signal-peptide
 FT Protein 33..395
 FT /note= "Human mature DME-5 protein"
 FT Peptide 1..35
 FT /label= Signal-peptide
 FT Protein 36..395
 FT /note= "Human mature DME-5 protein"
 FT
 XX WO200179468-A2.
 XX
 XX 25-OCT-2001.
 XX
 XX 12-APR-2001; 2001WO-US11869.
 XX
 XX 13-APR-2000; 2000US-197580P.
 XX 19-APR-2000; 2000US-198403P.
 XX 28-APR-2000; 2000US-200185P.
 XX 05-MAY-2000; 2000US-202234P.
 XX 11-MAY-2000; 2000US-203509P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Pollicky JL, Hafalia A, Burford N, Ring HZ, Lal P, Tribouley CM;
 XX Yao MG, Yue H, Tang YT, Patterson C, Das D, Sanjanwala MS;
 XX Gandhi AR, Reddy R, Khan FA, Baughn MR, Ramkumar J, Griffin JA;
 XX Au-Young J;
 XX
 XX WPI; 2002-066363/09.
 XX N-PDB; AAD4670.

PA (AKAM/) AKAMA T O.
 XX Fukuda MN, Akama TO;
 XX WPI; 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy
 PS Example 5, Page 24-25; 69pp; English.
 XX
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents mouse intestinal N-acetylglucosamine-6-
 CC sulfotransferase, which is given in comparison with (I) in the
 CC exemplification of the present invention.
 CC
 SQ Sequence 418 AA;
 QY
 ID 72 SMRSGSFTVGQLFSGHPDVFYLMFPAWHV 100
 DB
 RESULT 15
 ABB81559
 ID ABB81559 standard; Protein: 169 AA.
 XX
 AC ABB81559;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Human intestinal N-acetylglucosamine-6-sulfotransferase portion SEQ:7.
 XX
 KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme: GlcNAc6ST;
 KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.
 XX
 OS Homo sapiens.
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 09-AUG-2001; 2001US-0927602.
 XX
 PR 11-AUG-2000; 2000US-325773P.
 XX
 PA (FUKU/) FUKUDA M N.
 PA (AKAM/) AKAMA T O.
 XX
 PI Fukuda MN, Akama TO;
 XX
 DR WPI; 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy
 PS Example 2; Fig 3C; 69pp; English.
 XX
 CC The present invention describes human corneal

CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human intestinal
 CC N-acetylglucosamine-6-sulfotransferase (GlcNAc6ST), which is given in
 CC the exemplification of the present invention.
 CC
 SQ Sequence 169 AA;
 QY
 ID 16 SMRSGSFTVGQLFSGHPDVFYLMFPAWHV 44
 DB
 RESULT 16
 AAY72639
 ID AAY72639 standard; Protein: 390 AA.
 XX
 AC AAY72639;
 XX
 DT 02-MAY-2001 (first entry)
 XX
 DE Human glycosyl sulfotransferase-4alpha (GST-4alpha).
 XX
 KW Human: glycosyl sulfotransferase-4alpha; GST-4alpha; immunosuppressive;
 KW therapy; selectin binding inhibitor; gene therapy; inflammation;
 KW systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
 KW polyarteritis nodosa; polymyositis; systemic sclerosis; dermatitis;
 KW glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenailitis;
 KW Hashimoto's disease; Grave's disease; hypoparathyroidism; anemia;
 KW demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
 KW myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
 KW asthma; hypersensitivity; rheumatic fever; tissue rejection;
 KW chromosome 16q23.1.
 XX
 OS Homo sapiens.
 XX
 PN WO200106015-A1.
 XX
 PD 25-JAN-2001.
 XX
 PF 19-JUL-2000; 2000WO-0519741.
 XX
 PR 20-JUL-1999; 99US-0144694.
 XX
 PR 13-JUL-2000; 2000US-0593828.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Rosen SD, Lee JK, Hemmerich S;
 XX
 DR WPI; 2001-138471/14.
 DR N-PDB; AAD02697, AAD02698, AAD02699.
 XX
 PT New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
 PT diagnostic and therapeutic agent screening applications
 XX
 PS Claim 3; Fig 1; 128pp; English.
 XX
 CC The present sequence is human glycosyl sulfotransferase-4alpha (GST-4
 CC alpha). GST-4 gene is found on chromosome 16q23.1.
 CC GST is a type 2 membrane protein useful for inhibiting a binding event
 CC between a selectin and a selectin ligand, which comprises contacting the
 CC selectin with a non-sulphated selectin ligand, GST and a small molecular
 CC agent that inhibits the sulphation activity of GST. GST is also useful
 CC in inhibiting a selectin mediated binding event. GST is useful in gene

CC therapy to treat disorders such as acute or chronic inflammation,
CC systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
CC nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
CC glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
CC disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
CC anaemia, demyelinating diseases, cirrhosis, ulcerative colitis,
CC dermatitis, myocarditis, regional enteritis, adult respiratory distress
CC syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
CC bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
CC during transplantation.
CC Note: The present sequence is also shown in sequence listing (page
CC no: 56) but lacks four nucleotides at its 3' end.
XX
SQ Sequence 390 AA;
XX
Query Match 94.6%; Score 157; DB 22; Length 390;
Best Local Similarity 93.1%; Pred. No. 5.1e-16;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMRSGSSFVGQLEFGHDPVFLMEPAWHV 29
DB 49 SMRSGSSFVGQLEFGHDPVFLMEPAWHV 77
RESULT 17
ABB81556
ID ABB81556 standard; Protein; 390 AA.
XX
AC ABB81556;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human intestinal N-acetylglucosamine-6-sulfotransferase SEQ ID NO:4.
XX
KM Human; N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KM ophthalmological.
XX
OS Homo sapiens.
XX
PN US2002061562-A1.
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU/) FUKUDA M. N.
PA (AKAM/) AKAMA T. O.
XX
PI Fukuda MN, Akama TO;
XX
DR WPI: 2002-507643/54.
XX
PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy -
XX
PS Example 5; Fig 2A-B; 69pp; English.
XX
CC The present invention describes human corneal
CC N-acetylglucosamine-6-sulfotransferase (1), which is able to catalyse
CC sulfation of keratan sulfate (KS). Also described is a method for
CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC and detecting susceptibility to MCD. (1) is located to chromosome 16q22,
CC and has ophthalmological activity. (1) can be used to treat or prevent
CC macular corneal dystrophy types I or II. (1) makes possible treatment
CC of MCD without requiring keratoplasty or keratectomy. The present
CC sequence represents human intestinal N-acetylglucosamine-6-
CC sulfotransferase, which is given in comparison with (1) in the
CC exemplification of the present invention.

SQ Sequence 390 AA;
XX
Query Match 94.6%; Score 157; DB 23; Length 390;
Best Local Similarity 93.1%; Pred. No. 5.1e-16;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 SMRSGSSFVGQLEFGHDPVFLMEPAWHV 29
DB 49 SMRSGSSFVGQLEFGHDPVFLMEPAWHV 77
RESULT 18
ABB41947
ID ABB41947 standard; Protein; 418 AA.
XX
AC ABB41947;
XX
DT 08-FEB-2001 (first entry)
XX
DE Human OREF1711 polypeptide sequence SEQ ID NO:3422.
XX
KM Human; open reading frame; OREF; detection; cytosolic; hepatotropic;
KM vulnery; antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
KM anticonvulsant; osteopathic; antiarthritis; immunosuppressant; cardiant;
KM immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
KM hypotensive; dermatological; immunosuppressive; antiinflammatory;
KM antiviral; antibacterial; antifungal; antineumatic; antithyroid;
KM antianaemic; gene therapy; cancer; proliferative disorder; hypertension;
KM neurodegenerative disorder; osteoarthritis; graft vs host disease;
KM cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
KM cholesterol ester storage; systemic lupus erythematosus; infection;
KM severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
KM allergy; aplastic anaemia; nocturnal hemoglobinuria; burn; wound;
KM bone damage; cartilage damage; antiinflammatory disease; coagulation;
KM thrombosis; contraceptive.
XX
OS Homo sapiens.
XX
PN WO200058473-A2.
PD 05-OCT-2000.
XX
PF 31-MAR-2000; 2000WO-US08621.
XX
PR 31-MAR-1999; 99US-0127607.
PR 02-APR-1999; 99US-0127636.
PR 05-APR-1999; 99US-0127728.
PR 30-MAR-2000; 2000US-0540763.
XX
PA (CORA-) CORAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI: 2000-602362/57.
XX
PS N-PSDB: AAC76156.
XX
PT Novel nucleic acids and peptides derived from open reading frame X,
PT useful for treating e.g. cancers, proliferative disorders,
PT neurodegenerative disorders and cardiovascular disease -
XX
PS Claim 11; Page 2599-2600; 5507pp; English.
XX
CC AAC74446 to AAC77606 encode the proteins given in ABB40237 to ABB43397,
CC which represent the human OREF open reading frames 1 to 3161. The OREF
CC sequences have activities such as: cytosolic; hepatotropic; vulnery;
CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;
CC osteopathic; anticonvulsant; antiarthritis; immunosuppressant;
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;
CC antidiabetic; hypotensive; dermatological; immunosuppressive;
CC antiinflammatory; antibacterial; antifungal; antineumatic;
CC antithyroid; and antianaemic. The sequences can be used for determining
CC the presence of or predisposition to, or preventing or treating
CC pathological conditions associated with an OREF-associated disorder. The

CC nucleic acids can be used to express ORX proteins in gene therapy.
CC vectors. The proteins and nucleic acids may be used to treat cancers,
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,
CC graft vs host disease, cardiovascular disease, diabetes mellitus,
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,
CC nocturnal haemoglobinuria, antiinflammatory disease, to enhance
CC coagulation, to inhibit thrombosis, and as a contraceptive.

```

Query Match          94.6%   Score 157;  DB 21;   Length 418;
Best Local Similarity 93.1%   Pred. No. 5.5e-16;
Matches 27;  Conservative 1;  Mismatches 1;  Indels 0;  Gaps 0

```

```

QY      1 SWRSGSSFFVQLFGQHPDVFYIMEPAWHV 29
          |||||:|||||
Db      76 SWRSGSSFLGLFSQHPDVFYIMEPAWHV 104

```

RESULT 19
ABB81561
ID ABB81561 standard; Protein; 169 AA

Query Match	74.7%	Score 124;	DB 23;	Length 169;
Best Local Similarity	69.0%	Pred. No. 2.7e-11;		
Matches 20;	Conservative 5;	Mismatches 4;	Indels 0;	Gaps 0

```

Oy      1  SWRSGSSFYVGLFGQHPDVFYLMPEAWHY 29
          :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      17  TWRSGSSFFGELFNQNPVEVFFLYEPVWHY 45

```

RESULT 20
AAV31656
ID AAV31656 standard; Protein; 483 AA

```

Query Match      74.7%   Score 124   DB 20:   Length 483;
Best Local Similarity 69.0%   Pred. No. 9,2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY      1  SWRGSSEFVQLTGQHPDYFLMEPRMAY 29
      : : : : : : : : : : : : : : : :
Db      125  TWRGSSFEFGELNQNPRVEFLDEPMVHY 153

```

```

QY      1  SWRSGSSFVGOLFQGHDPVFYIMEPAWHY 29
      :|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      125 TWRRSGSSFEGELFNQNPVEFFLYEPVWHY 153

```

```

RESULT 21
AAAY31657
XX AAY31657 standard; Protein; 484 AA.
AC
XX AAY31657;
AC
XX 09-NOV-1999 (first entry)
DE Human N-acetylglucosamine-6-O-sulfotransferase.
XX
XX N-acetylglucosamine-6-O-sulfotransferase; human; GLYCAM-1;
KM L-selectin ligand.
XX
XX Homo sapiens.
OS
XX EP943688-A2.
XX
XX 22-SEP-1999.
XX
XX 04-MAR-1999; 99EP-0301530.
XX
XX 24-JUN-1998; 98JP-0177844.
XX
XX 05-MAR-1998; 98JP-0054007.
XX
XX (SEK) SEIKAGAKU CORP.
XX
XX Habuchi O, Kadamatsu K, Kanagaki R, Muramatsu H;
PI Muramatsu T, Uchimura K;
XX
XX MPI; 1999-520337/44.
XX
XX N-PSDB; AAX87821.
XX
XX New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PT for synthesis of sugar chains, e.g. GLYCAM-1
XX
XX Claim 2; Page 28-30; 41pp; English.
XX
XX The present sequence represents human N-acetylglucosamine-6-O-
CC sulfotransferase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC beta 1-3 = beta 1-3 glycosidic linkage, and beta 1-4 = beta 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GLYCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87821) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GLYCAM-1 using transformants which harbour
CC the DNA.
XX
XX
SQ Sequence 484 AA;
XX
XX Query Match 74.7%; Score 124; DB 20; Length 484;
XX Best Local Similarity 69.0%; Pred. No. 9.2e-11;
XX Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SWRSSSFVGQFGHPDYFLMEPAHV 29
Db 126 TWRSSSFVGELFNQNPVEFLYEPVHW 154

```

```

XX
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300252.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX MPI; 2001-318749/34.
XX
XX
XX Claim 8; SEQ ID 17679; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification. Where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX
SQ Sequence 530 AA;
XX
XX Query Match 74.7%; Score 124; DB 22; Length 530;
XX Best Local Similarity 69.0%; Pred. No. 1e-10;
XX Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SWRSSSFVGQFGHPDYFLMEPAHV 29
Db 172 TWRSSSFVGELFNQNPVEFLYEPVHW 200

```

```

RESULT 22
AAB95367
XX AAB95367 standard; Protein; 530 AA.
XX
XX AAB95367;
AC
XX
XX 26-JUN-2001 (first entry)
DT
XX
XX Human protein sequence SEQ ID NO:17679.
DE

```

```

RESULT 23
AAU69414
XX AAU69414 standard; Protein; 531 AA.
XX
XX AAU69414;
AC
XX
XX 30-JAN-2002 (first entry)
DT
XX
XX

```


XX	(SEGR) SEIKAGAKU KOSYO CO LTD.
PA	
XX	
PI	Fukuta M, Habuchi O;
XX	
DR	NPI; 1998-268750/26.
XX	N-PDB; AAV36418.
PT	Keratan sulphate 6-sulpho-transferase polypeptide - transfers
PT	sulphate from sulphate donor to galactose 6-hydroxy group etc.
XX	
PS	ClaM 4; Page 15-16; 21pp; English.
XX	
CC	This sequence produces the protein KSGal6ST (Keratan sulphate
CC	6-sulphotransferase), which has a molecular weight of about 46700. The
CC	cDNA sequence of this protein was obtained by radiolabelling the cDNA of
CC	chick chondroitin 6-sulphotransferase (C6ST) and using this as a probe in
CC	a random oligonucleotide-primed labelling method. Human foetal brain
CC	cDNA was inserted into a phage lambda gfil cloning vector whereby the
CC	clones containing the KSGal6ST were obtained by hybridization using the
CC	prepared probe. The positive clones were subcloned into a recombinant
CC	expression vector and used to transform COS-7 cells, from which cells
CC	expressing KSGal6ST can be selected. The phage cDNA inserts were
CC	isolated and subcloned into a Bluescript plasmid. Deletion clones were
CC	then prepared from which both strands were sequenced by the Sanger
CC	method. The KSGal6ST of the invention transfers the sulphate from a
CC	sulphate donor to galactose 6-OH groups in keratan sulphate, but does not
CC	transfer sulphate to chondroitin, chondroitin sulphate A or C, dermatan
CC	sulphate or CDNS heparin.
SQ	Sequence 411 AA;
Query Match	69.9%; Score 116; DB 19; Length 411;
Best Local Similarity	81.5%; Pred. No. 1.4e-09;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;	
Yq	3 RSGSSEFGQLFQGHPDYFIEMPMHV 29
	:
Db	70 RSGSSEFGQLFNCHLDVFLFEPLVHV 96
RESULT 26	
ABB81563	
ID	ABB81563 standard; Protein; 174 AA.
AC	ABB81563;
DT	05-SEP-2002 (first entry)
XX	
DE	Human chondroitin-6-sulfoyltransferase portion SEQ ID NO:11.
KW	Human; N-acetylglucosamine-6-sulfoyltransferase; enzyme; GlcNAc6ST;
RW	corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
XX	ophthalmology.
XX	
OS	Homo sapiens.
XX	
PN	US2002061562-A1.
PD	23-MAY-2002.
PX	09-AUG-2001; 2001US-0927602.
PF	11-AUG-2000; 2000US-325773P.
PR	
PA	(FUKU/) FUKUDA M N.
PA	(AKAM/) AKAMA T O.
XX	
PI	Fukuda MN, Akama TO;
XX	
DR	WPI; 2002-507643/54.
XX	
PT	New nucleic acid encoding corneal

PT	N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT	monitoring and diagnosis of macular corneal dystrophy -
XX	
PS	Example 2; Fig 3C; 69pp; English.
XX	
CC	The present invention describes human corneal
CC	N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyse
CC	sulfation of keratan sulfate (KS). Also described is a method for
CC	monitoring the effect of treatments for macular corneal dystrophy (MCD),
CC	and detecting susceptibility to MCD. (I) is located to chromosome 16q42,
CC	and has opticalmolecular activity. (II) can be used to treat or prevent
CC	macular corneal dystrophy types I or II. (I) makes possible treatment
CC	of MCD without requiring keratoplasty or keratectomy. The present
CC	sequence represents a portion of the human chondroitin-6-
CC	sulfotransferase, which is given in the exemplification of the present
CC	invention.
XX	
SQ	Sequence 1/4 AA:
Query Match	60.8%; Score 101; DB 23; Length 174;
Best Local Similarity	59.3%; Pred.No. 1.le-07;
Matches 16; Conservative	5; Mismatches 6; Indels 0; Gaps 0;
DY	3 RSGSFTVGOLFEGOHDPVFLYLMPEAMHV 29 1: :- -:: : 22 RTGSSFFVEGFENOGNIFLYLFEPLWHI 48
DB	
RESULT 27	
AAM06480	
ID AAM06480 standard; Protein; 458 AA.	
AC AAM06480;	
XX	
XX	AAW06480; (first entry)
DT XX	
DE Chick chondroitin 6-sulphotransferase.	
XX Chondroitin 6-sulphotransferase; C6ST; chondroitin sulphate; proteoglycan. XX KW XX OS Gallus sp. XX	
FH Key Location/Qualifiers FT Peptide 1..33 FT /label= Sig-peptide 34..458 FT Protein /label= Mat-protein /note= "preferred polypeptide (Claim 5)" FT FT 1..458 FT /note= "preferred polypeptide (Claim 3)" FT FT 20..458 FT /note= "preferred polypeptide (Claim 4)" FT FT 38..187 FT /note= "preferred polypeptide (Claim 6)" FT FT 24..37 FT /label= Transmembrane_domain 62..64 FT /label= Glycosylation /note= "potential N-glycosylation site" FT FT 73..75 FT /label= Glycosylation /note= "potential N-glycosylation site" FT FT 95..97 FT /label= Glycosylation /note= "potential N-glycosylation site" FT FT 236..238 FT /label= Glycosylation /note= "potential N-glycosylation site" FT FT 399..401 FT /label= Glycosylation /note= "potential N-glycosylation site" FT Modified-site 443..445	


```

FT      /label- Glycosylation
FT      /note= "potential N-glycosylation site"
XX
XX      EP745668-A2.
XX
XX      04-DEC-1996.
XX
XX      30-MAY-1996; 96EP-0303887.
XX
XX      31-MAY-1995; 95JP-0134358.
XX
XX      (SEK ) SEIKAGAKU KOGYO CO LTD.
XX
XX      Fukuta M, Habuchi O:
XX
XX      WPI; 1997-013692/02.
XX
XX      N-PSDB; AAT45037.
XX
XX      DNA encoding chondroitin 6-sulphotransferase - for recombinant
XX      production of C6ST, for use in industrial processes
XX
XX      Claim 2: Page 18-20; 30pp; English.
XX
XX      Chick embryo chondrocyte chondroitin 6-sulphotransferase (C6ST)
XX      (AAW06480) is capable of catalysing the transfer of a sulphate group
XX      from 3'-phosphoadenosine 5' phosphosulphate to the hydroxyl group
XX      at the C-6 position of the N-acetylgalactosamine residue of
XX      chondroitin. Its amino acid sequence was deduced from a cDNA clone
XX      (AAI45037) isolated from a chick embryo chondrocyte cDNA library.
XX      Isolation of the cDNA allows the mass prodn. of C6ST in transformed
XX      host cells. Recombinant C6ST polypeptides are useful for analysing
XX      the activities of chondroitin sulphate and for modifying its
XX      function, and for raising antibodies.
XX
XX      Sequence 458 AA:
XX
XX      Query Match 60.8%; Score 101; DB 18; Length 458;
XX      Best Local Similarity 59.3%; Pred. No. 3.4e-07;
XX      Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
XX
XX      QY 3 RSGSFFVGQIFGHPDVFYLMPEAMHV 29
XX      122 RTGSSFFVGEEFNQGNIFYLEPLWHI 148
XX
XX      RESULT 28
XX      AAW52863
XX      ID AAW52863 standard; Protein: 479 AA.
XX
XX      AC AAW52863;
XX
XX      DT 30-JUN-1998 (first entry)
XX
XX      DE Glycosaminoglycan sulphotransferase protein.
XX
XX      KW Glycosaminoglycan sulphotransferase; C6ST; N-acetylglactosamine;
XX      galactose; glycosaminoglycan; chondroitin; keratan sulphate; diagnostic.
XX
XX      OS Homo sapiens.
XX
XX      PN EP821066-A1.
XX
XX      PD 28-JAN-1998.
XX
XX      PE 22-JUL-1997; 97EP-0305476.
XX
XX      PR 24-JUL-1996; 96JP-0195063.
XX
XX      (SEK ) SEIKAGAKU KOGYO CO LTD.
XX
XX      Fukuta M, Habuchi O;
XX
XX      WPI; 1998-102622/10.

```

```

XX      Human glycosaminoglycan sulpho:transferase protein - catalyses
XX      6-sulphation of chondroitin to produce chondroitin sulphate for use
XX      in pharmaceutical(s)
XX
XX      Claim 3; Page 21-23; 27pp; English.
XX
XX      The human glycosaminoglycan sulphotransferase (C6ST) is a type II
XX      membrane protein derived from foetal brain tissue with a molecular weight
XX      of 50-55 kD. The protein transfers sulphate groups from a donor to the
XX      N-acetylgalactosamine residue or galactose residue of glycosaminoglycan.
XX      C6ST displays substrate specificity transferring the sulphate group to
XX      the hydroxyl group position at C-6 of the N-acetylgalactosamine residue
XX      of chondroitin and the hydroxyl group position at C-6 of the galactose
XX      residue of keratan sulphate. The protein is useful for studying the
XX      function of chondroitin sulphate and can provide chondroitin sulphate
XX      that may be useful in pharmaceuticals. Both the protein and DNA might
XX      be useful for treating or diagnosing diseases attributable to low C-6
XX      sulphation of chondroitin N-acetylgalactosamine residues.
XX
XX      Sequence 479 AA:
XX
XX      Query Match 60.8%; Score 101; DB 19; Length 479;
XX      Best Local Similarity 59.3%; Pred. No. 3.5e-07;
XX      Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
XX
XX      QY 3 RSGSFFVGQIFGHPDVFYLMPEAMHV 29
XX      142 RTGSSFFVGEEFNQGNIFYLEPLWHI 168
XX
XX      RESULT 29
XX      ABB64513
XX      ID ABB64513 standard; Protein: 315 AA.
XX
XX      AC ABB64513;
XX
XX      DT 26-MAR-2002 (first entry)
XX
XX      DE Drosophila melanogaster polypeptide SEQ ID NO 20331.
XX
XX      KW Drosophila; developmental biology; cell signalling; insecticide;
XX      pharmaceutical.
XX
XX      OS Drosophila melanogaster.
XX
XX      PN W0200171042-A2.
XX
XX      PD 27-SEP-2001.
XX
XX      PE 23-MAR-2001; 2001WO-US09231.
XX
XX      PR 23-MAR-2000; 2000US-191637P.
XX
XX      PR 11-JUL-2000; 2000US-0614150.
XX
XX      (PEKE ) PE CORP NY.
XX
XX      Venter JC, Adams M, Li PWD, Myers EW;
XX
XX      WPI; 2001-656860/75.
XX
XX      N-PSDB; ABL08616.
XX
XX      New isolated nucleic acid detection reagent for detecting 1000 or more
XX      genes from Drosophila and for elucidating cell signalling and cell-cell
XX      interactions -
XX
XX      Disclosure; SEQ ID NO 20331; 21pp + Sequence Listing; English.
XX
XX      The invention relates to an isolated nucleic acid detection reagent
XX      capable of detecting 1000 or more genes from Drosophila. The invention is
XX      useful in developmental biology and in elucidating cell signalling and
XX      cell-cell interactions in higher eukaryotes for the development of
XX      insecticides, therapeutics and pharmaceutical drugs. The invention

```


DT 09-AUG-2002 (first entry)

XX Human transporter and ion channel (TRICH) 16.

DE Human transporter and ion channel (TRICH) 16.

KM Human: transporter and ion channel; TRICH; transport disorder;

KM diabetes mellitus; angina; Alzheimer's disease; neurological disorder;

KM stroke; Huntington's disease; meningitis; muscle; myocarditis; cancer;

KM infectious myositis; arrhythmia; asthma; immunological; gene therapy;

KM acquired immunodeficiency syndrome; AIDS; allergy; atherosclerosis;

KM cell proliferative disorder; cerebroprotective; cirrhosis; hepatitis;

KM transgenic; neuroprotective; anticonvulsant; nootropic; cytostatic;

KM antiinflammatory; hepatotropic; psoriasis.

OS Homo sapiens.

XX Location/Qualifiers

XX Key 177..193

XX Domain /label=Transmembrane_domain

XX Domain 348..366

XX Domain /label=Transmembrane_domain

XX Modified-site 393..429

XX Modified-site /note="ATPase phosphorylation site"

XX Modified-site 417..471

XX Modified-site /note="ATPase phosphorylation site"

XX Modified-site 431..449

XX Modified-site /note="ATPase phosphorylation site"

XX Modified-site 575..585

XX Modified-site /note="ATPase phosphorylation site"

XX Modified-site 644..684

XX Modified-site /note="ATPase phosphorylation site"

XX WO200222684-A2.

XX 21-MAR-2002.

XX 14-SEP-2001; 2001WO-US28938.

XX 15-SEP-2000; 2000US-232685P.

XX 22-SEP-2000; 2000US-234842P.

XX 29-SEP-2000; 2000US-236882P.

XX 05-OCT-2000; 2000US-238057P.

XX 13-OCT-2000; 2000US-240540P.

XX 18-OCT-2000; 2000US-241700P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lee EA, Yue H, Lai PC, Walla NK, Baughn MR, Warren BA, Lee S,

XX Sanjanaala MS, Yao MG, Ramkumar J, Thornton M, Gandhi AR, Naini A,

XX Policky JL, Elliott VS, Arvizu C, Raumann BE, Bruns CM, Naini A,

XX Hatfield AA, Nguyen DB, Xu Y, Lu DM, Ison CH, Griffin JA,

XX Reddy RM, Burford N;

XX WPI; 2002-393948/42.

XX N-PSDB; AAD36313.

XX Polypeptides of human transporters and ion channels, useful for

XX diagnosing, treating or preventing transport, neurological, muscle,

XX immunological and cell proliferative disorders

XX Claim 1; Page 163-165; 204pp; English.

XX The invention relates to human transporters and ion channels (TRICH)

XX and their corresponding nucleic acid sequences. TRICH is useful for

XX screening an agonist/antagonist that modulates its activity. TRICH is

XX useful as an immunogen for preparing antibodies which are useful for

XX diagnosing a condition of disease associated with its expression in a

XX subject, and for detecting and purifying it from a sample. TRICH DNA

XX is useful as probe or a primer for assessing toxicity of a test

XX compound. Composition comprising TRICH or its agonist is useful for

XX treating a disease or condition associated with decreased expression

XX of functional TRICH and composition comprising TRICH antagonist is

XX useful for treating a disease or condition associated with TRICH

XX overexpression of TRICH. TRICH sequence is used in the diagnosis and

CC treatment of transport disorder e.g. diabetes mellitus, angina,

CC Alzheimer's disease; neurological disorder e.g. epilepsy, stroke,

CC Huntington's disease; bacterial and viral meningitis; muscle disorder

CC e.g. myocarditis, infectious myositis, arrhythmias, asthma,

CC immunological disorder e.g. acquired immunodeficiency syndrome (AIDS),

CC allergies, atherosclerosis; and cell proliferative disorders e.g.

CC cirrhosis, hepatitis, psoriasis and cancers. TRICH DNA is used in

CC gene therapy. TRICH DNA is useful for creating knockin humanised

CC animals (pigs) or transgenic animals (mice or rats) to model human

CC disease. The present sequence is human TRICH protein.

XX

XX Sequence 791 AA:

XX

XX Query Match 30.1%; Score 50; DB 23; Length 791;

XX Best Local Similarity 53.8%; Pred. No. 58;

XX Matches 14; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

XX

XX 4 SGSSFVGQLFGQHPDVFYLMPPAHV 29

XX :||| |||||

XX Db 37 AGSIF--SFGILPLVVFY-WRPAMHV 58

XX

XX RESULT 33

XX ID ABB57815

XX ABB57815 standard; Protein: 4097 AA.

XX AC ABB57815;

XX 26-MAR-2002 (first entry)

XX

XX Drosophila melanogaster polypeptide SEQ ID NO 237.

XX

XX Drosophila; developmental biology; cell signalling; insecticide;

XX pharmaceutical.

XX Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US09231.

XX 23-MAR-2000; 2000US-194637P.

XX 11-JUL-2000; 2000US-0644150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

XX N-PSDB; ABL01918.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

XX genes from Drosophila and for elucidating cell signalling and cell-cell

XX interactions

XX Disclosure; SEQ ID NO 237; 21pp + Sequence listing; English.

XX

XX The invention relates to an isolated nucleic acid detection reagent

XX capable of detecting 1000 or more genes from Drosophila. The invention is

XX useful in developmental biology and in elucidating cell signalling and

XX cell-cell interactions in higher eukaryotes for the development of

XX insecticides, therapeutics and pharmaceutical drugs. The invention

XX discloses genomic DNA sequences (AB116175-AB130511), expressed DNA

XX sequences (AB101840-AB116175) and the encoded proteins

XX (ABB57737-ABB72072).

XX The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pcl_sequences.

XX Sequence 4097 AA;

XX Homo sapiens.
 OS
 PN WO200200855-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 26-JUN-2001; 2001WO-EP07297.
 XX
 PR 26-JUN-2000; 2000US-214012P.
 PR 14-DEC-2000; 2000US-255148P.
 PA (FARB) BAYER AG.
 XX
 PI xiao Y;
 XX
 DR WPI: 2002-130886/17.
 DR N-PDB; ABR13701.
 PT
 PT New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
 PT useful for identifying modulators of enzyme activity for treating
 PT cardiovascular disease, diabetes, obesity and hyperlipidaemia
 PS
 PS Claim 1; Fig 2; 123pp; English.
 XX
 CC The invention describes a purified human S-acyl fatty acid synthase
 CC thioesterase (SFS1)-like enzyme (I). (I) is useful for screening for
 CC agents which decrease or regulate the activity of an SFS1-like enzyme and
 CC agents which decrease extracellular matrix degradation. The
 CC polynucleotide (II) is useful for detecting a polynucleotide which
 CC encodes (I) in a biological sample by formation of a hybridisation
 CC complex. A reagent modulating the activity of SFS1-like enzyme or an
 CC antibody can also be used to detect the polypeptide or polynucleotide in
 CC a biological sample. A reagent modulating the activity of (I) or (II) is
 CC also useful for treating a SFS1-like enzyme dysfunction related disease
 CC condition such as cardiovascular disease, hyperlipidaemia, obesity,
 CC anorexia, cachexia, wasting disorders, appetite suppression, appetite
 CC enhancement, bulimia or diabetes. The cardiovascular diseases treated by
 CC the above mentioned methods and reagents include congestive heart
 CC failure, myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases, and
 CC peripheral vascular diseases. (I) is useful in diagnostic assays for
 CC detecting diseases and abnormalities related to presence of mutations in
 CC the nucleic acid sequences which encode the enzyme. This is the amino
 CC acid sequence of a human s-acyl fatty acid synthase thioesterase-like
 CC enzyme, described in the method of the invention.
 CC
 SQ Sequence 265 AA:
 XX
 Query Match 29.5%; Score 49; DB 23; Length 265;
 Best Local Similarity 41.4%; Pred. No. 23;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 OY 1 SMR---SGSSFVQGLFGQHPDVTYIMEPA 26
 DB 219 AMKDVTSGNAKIYQLPGH---FYLLDPA 244
 RESULT 38
 AAU75504
 ID AAU75504 standard; Protein; 265 AA.
 AC AAU75504;
 XX
 DT 23-APR-2002 (first entry)
 XX
 DE S-acyl fatty acid synthase thioesterase-like enzyme (short).
 XX
 XX S-acyl fatty acid synthase thioesterase-like; SFS1-like; enzyme;
 KM cardiac; anorectic; vasotropic; extracellular matrix degradation;
 KM cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;
 KM wasting disorder; appetite suppression; appetite enhancement; bulimia;
 KM diabetes; congestive heart failure; myocardial infarction; human;

KM ischaemic disease; atrial arrhythmia; ventricular arrhythmia;
 KM hypertensive vascular disease; peripheral vascular disease.
 XX
 OS Homo sapiens.
 PN WO200200855-A2.
 XX
 PD 03-JAN-2002.
 XX
 PF 26-JUN-2001; 2001WO-EP07297.
 XX
 PR 26-JUN-2000; 2000US-214012P.
 PR 14-DEC-2000; 2000US-255148P.
 PA (FARB) BAYER AG.
 XX
 PI xiao Y;
 XX
 DR WPI: 2002-130886/17.
 DR
 PT
 PT New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
 PT useful for identifying modulators of enzyme activity for treating
 PT cardiovascular disease, diabetes, obesity and hyperlipidaemia
 PS
 PS Disclosure: Fig 14; 123pp; English.
 XX
 CC The invention describes a purified human S-acyl fatty acid synthase
 CC thioesterase (SFS1)-like enzyme (I). (I) is useful for screening for
 CC agents which decrease or regulate the activity of an SFS1-like enzyme and
 CC agents which decrease extracellular matrix degradation. The
 CC polynucleotide (II) is useful for detecting a polynucleotide which
 CC encodes (I) in a biological sample by formation of a hybridisation
 CC complex. A reagent modulating the activity of SFS1-like enzyme or an
 CC antibody can also be used to detect the polypeptide or polynucleotide in
 CC a biological sample. A reagent modulating the activity of (I) or (II) is
 CC also useful for treating a SFS1-like enzyme dysfunction related disease
 CC condition such as cardiovascular disease, hyperlipidaemia, obesity,
 CC anorexia, cachexia, wasting disorders, appetite suppression, appetite
 CC enhancement, bulimia or diabetes. The cardiovascular diseases treated by
 CC the above mentioned methods and reagents include congestive heart
 CC failure, myocardial infarction, ischemic diseases of the heart, atrial
 CC and ventricular arrhythmia, hypertensive vascular diseases, and
 CC peripheral vascular diseases. (I) is useful in diagnostic assays for
 CC detecting diseases and abnormalities related to presence of mutations in
 CC the nucleic acid sequences which encode the enzyme. This sequence encodes
 CC a human s-acyl fatty acid synthase thioesterase-like enzyme, described
 CC in the method of the invention.
 CC
 SQ Sequence 265 AA:
 XX
 Query Match 29.5%; Score 49; DB 23; Length 265;
 Best Local Similarity 41.4%; Pred. No. 23;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
 OY 1 SMR---SGSSFVQGLFGQHPDVTYIMEPA 26
 DB 219 AMKDVTSGNAKIYQLPGH---FYLLDPA 244
 RESULT 39
 AAB93499
 ID AAB93499 standard; Protein; 318 AA.
 AC AAB93499;
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12816.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KM
 OS Homo sapiens.

PN EPI074617-A2.
XX
XX 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
PR 27-AUG-1999; 99JP-0300253.
PR 11-JAN-2000; 2000JP-0118776.
PR 02-MAY-2000; 2000JP-0183767.
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 12816; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
CC full-length cDNAs defined in the specification, where a primer set
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises at least 15 nucleotides and the combination
CC of oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.
XX
XX Sequence 318 AA:
SQ
Query Match 29.5%; Score 49; DB 22; Length 318;
Best Local Similarity 41.4%; Pred. No. 29;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
QY 1 SMR---SGSFFVGOLFQGPDPVFYIEMEA 26
Db 272 AWKDVTSGNAKIYOLPGH---FYLLDPA 297
:|: ||::: || | ||:::|
AAU75505
ID AAU75505 standard; Protein; 318 AA.
XX
XX AAU75505;
AC
XX
XX 23-APR-2002 (first entry)
DT
XX
XX S-acyl fatty acid synthase thioesterase-like enzyme (long).
DE
XX
XX S-acyl fatty acid synthase thioesterase-like; SFST-like; enzyme;
KW cardiac; anorectic; vasotropic; extracellular matrix degradation;
KM cardiovascular disease; hyperlipidaemia; obesity; anorexia; cachexia;

KM wasting disorder; appetite suppression; appetite enhancement; bulimia;
KM diabetes; congestive heart failure; myocardial infarction; human;
KM ischaemic disease; atrial arrhythmia; ventricular arrhythmia;
KM hypertensive vascular disease; peripheral vascular disease.
XX
XX Homo sapiens.
OS
XX
XX WO200200855-A2.
PN
XX
XX 03-JAN-2002.
PD
XX
XX 26-JUN-2001; 2001WO-EP07297.
PF
XX
XX 26-JUN-2000; 2000US-214012P.
PR 14-DEC-2000; 2000US-255148P.
XX
XX (FARB) BAYER AG.
PA
XX
XX Xiao Y;
PI
XX
XX WPI: 2002-130886/17.
DR
XX
XX N-PSDB; ABK13711.
DR
XX
XX New purified human S-acyl fatty acid synthase thioesterase-like enzyme,
PT useful for identifying modulators of enzyme activity for treating
PT cardiovascular disease, diabetes, obesity and hyperlipidaemia -
XX
XX Disclosure; Fig 16; 123pp; English.
XX
XX The invention describes a purified human S-acyl fatty acid synthase
CC thioesterase (SFST)-like enzyme (I). (I) is useful for screening for
CC agents which decrease or regulate the activity of an SFST-like enzyme and
CC agents which decrease extracellular matrix degradation. The
CC polynucleotide (II) is useful for detecting a polynucleotide which
CC encodes (I) in a biological sample by formation of a hybridisation
CC complex. A reagent modulating the activity of SFST-like enzyme or an
CC antibody can also be used to detect the polypeptide or polynucleotide in
CC a biological sample. A reagent modulating the activity of (I) or (II) is
CC also useful for treating a SFST-like enzyme dysfunction related disease
CC condition such as cardiovascular disease, hyperlipidaemia, obesity,
CC anorexia, cachexia, wasting disorders, appetite suppression, appetite
CC enhancement, bulimia or diabetes. The cardiovascular diseases treated by
CC the above mentioned methods and reagents include congestive heart
CC failure, myocardial infarction, ischaemic diseases of the heart, atrial
CC and ventricular arrhythmia, hypertensive vascular diseases, and
CC peripheral vascular diseases. (I) is useful in diagnostic assays for
CC detecting diseases and abnormalities related to presence of mutations in
CC the nucleic acid sequences which encode the enzyme. This sequence encodes
CC a human S-acyl fatty acid synthase thioesterase-like enzyme, described
CC in the method of the invention.
XX
XX Sequence 318 AA:
SQ
Query Match 29.5%; Score 49; DB 23; Length 318;
Best Local Similarity 41.4%; Pred. No. 29;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;
QY 1 SMR---SGSFFVGOLFQGPDPVFYIEMEA 26
Db 272 AWKDVTSGNAKIYOLPGH---FYLLDPA 297
:|: ||::: || | ||:::|
AAU73357
ID AAU73357 standard; protein; 931 AA.
XX
XX AAU73357;
AC
XX
XX 15-NOV-1999 (first entry)
DT
XX
XX Group B Streptococcus (GBS) antigen (clone 4).
DE
XX
XX Group B Streptococcus; GBS; antigen; vaccine; Streptococcus infection;
KW

KM sepsis; meningitis; pneumonia; immunocompromised; diabetes; liver disease;
 XX cancer; veterinary; mastitis.
 XX
 OS Streptococcus sp.
 XX
 FH Key Location/Qualifiers
 FT MISC-difference 301
 FT /Label= unknown
 FT /note= "encoded by CCY"
 XX
 PN WO9942588-A2.
 XX
 PD 26-AUG-1999.
 XX
 PF 17-FEB-1999; 99WO-CA00114.
 XX
 PR 20-FEB-1998; 98US-0075425.
 XX
 PA (BIOC-) BIOCHEM VACCINS INC.
 XX
 PI Boyer M, Brodeur BR, Charlebois I, Hamel J, Martin D;
 PI Rioux C;
 XX
 DR WPI: 1999-540309/45.
 DR N-PSDB; AAA91106.
 XX
 PT Novel group B Streptococcus antigens - useful as vaccine
 PT compositions for prophylaxis or therapy of Streptococcus infections
 XX
 PS Claim 26; Fig 4E; 154pp; English.
 XX
 CC The invention provides Group B Streptococcus (GBS) antigens
 CC (AA927336-370) and nucleic acids (AA91103-X91111) encoding the antigens.
 CC The GBS antigens can be recombinantly expressed using standard
 CC recombinant methodology. The GBS antigens of the invention can be used as
 CC vaccine components for the treatment or prophylaxis of diseases and
 CC symptoms mediated by Streptococcus infection, especially group A
 CC Streptococcus (S. pyogenes), GBS or S. agalactiae, S. dysgalactiae, S.
 CC uberis, S. noxia, as well as staphylococcus aureus. The vaccines are
 CC administered to those individuals at risk of GBS infection, particularly
 CC pregnant women and infants for sepsis, meningitis, and pneumonia, as well
 CC as immunocompromised individuals, such as those with diabetes, liver
 CC disease or cancer. The vaccines also have veterinary applications, such as
 CC for the treatment of mastitis in cattle. The present sequence represents
 CC a GBS antigen of the invention.
 XX
 SQ Sequence 931 AA;
 Query Match 29.2%; Score 48.5; DB 20; Length 931;
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;
 Matches 10; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
 QY 1 SWRSSSFVGO-LFGQHPDV 19
 DB 304 TWPDSNFMVQGLYGRYDV 323
 RESULT 42
 ID AAY91279
 XX AAY91279 standard; Protein: 1250 AA.
 AC AAY91279;
 XX
 DT 30-MAY-2000 (first entry)
 XX
 DE Group B Streptococcus protein sequence SEQ ID NO:5.
 XX
 KW Group B Streptococcus; Streptococcus agalactiae; protein antigen;
 KW vaccine; screening; immunogen; detection; diagnosis; infection;
 KW antibody; antibody; antibacterial.
 XX
 OS Streptococcus agalactiae.
 XX

PN WO200006736-A2.
 XX
 PD 10-FEB-2000.
 XX
 PF 27-JUL-1999; 99WO-GB02444.
 XX
 PR 27-JUL-1998; 98GB-0016335.
 PR 19-MAR-1999; 99US-0125163.
 XX
 PA (MICR-) MICROBIAL TECHNIQS LTD.
 XX
 PI Le Page RWF, Wells JM, Hanniffy SB;
 PI
 DR WPI: 2000-195299/17.
 XX
 PT New Group B Streptococcus protein, useful as vaccine, for diagnosis of
 PT Streptococcal infections and for screening of antibodies or antibodies
 XX
 PS Claim 1; Fig 1; 123pp; English.
 XX
 CC AAA05803 to AAA05872 encode proteins, polypeptides and peptides (given
 CC in AAY91275 to AAY91343) isolated from Group B Streptococcus (GBS), also
 CC known as Streptococcus agalactiae. The GBS polynucleotides and
 CC polypeptides have antibacterial activity. Immunogenic compositions
 CC comprising GBS polynucleotides or polypeptides can be used as vaccines
 CC and for the treatment or prophylaxis of GBS infection. The
 CC polynucleotides and polypeptides can also be used in the detection of GBS
 CC and for screening DNA encoding bacterial cell envelope associated or
 CC secreted antigens in gram positive bacteria. AAA05873 to AAA05941
 CC represent primers used in the exemplification of the present invention.
 XX
 SQ Sequence 1250 AA;
 Query Match 29.2%; Score 48.5; DB 21; Length 1250;
 Best Local Similarity 50.0%; Pred. No. 1.7e+02;
 Matches 10; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
 QY 1 SWRSSSFVGO-LFGQHPDV 19
 DB 304 TWPDSNFMVQGLYGRYDV 323
 RESULT 43
 ID ABP26711
 XX ABP26711 standard; Protein: 1252 AA.
 AC ABP26711;
 XX
 DT 02-JUL-2002 (first entry)
 XX
 DE Streptococcus polypeptide SEQ ID NO 2598.
 XX
 KW Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
 KW group A streptococcus; Streptococcus pyogenes; antibacterial;
 KW antiinflammatory; infection; vaccine; meningitis; gene therapy.
 XX
 OS Streptococcus agalactiae.
 XX
 PN WO200234771-A2.
 XX
 PD 02-MAY-2002.
 XX
 PF 29-OCT-2001; 2001WO-GB04789.
 XX
 PR 27-OCT-2000; 2000GB-0026333.
 PR 24-NOV-2000; 2000GB-0028727.
 PR 07-MAR-2001; 2001GB-0005640.
 XX
 PA (CHIR-) CHIRON SPA.
 PA (GENO-) INST GENOMIC RES.
 XX
 PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
 PI Tectelin H;
 PI

XX DR WPI: 2002-352536/38.
XX DR N-PSDB: ABN67342.
XX PT New Streptococcus protein for the treatment or prevention of infection
XX PT or disease caused by Streptococcus bacteria, such as meningitis, and
XX PT for detecting a compound that binds to the protein -
XX PS Claim 1; Page 3409; 4525pp; English.
XX XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
XX antibodies that bind (1) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (1) are used to detect Streptococcus in a
XX biological sample. (1) is used to determine whether a compound binds to
XX (1). A composition comprising (1) or a nucleic acid encoding (1), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (1) may be used to recombinantly produce (1) and may be
XX used in gene therapy. Antibodies to (1) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX SQ Sequence 1252 AA;
XX
XX Query Match 29.2%; Score 48.5; DB 23; Length 1252;
XX Best Local Similarity 50.0%; Pred. No. 1.7e+02;
XX Matches 10; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
OY 1 SWRSGSSFVQG-LFGQHPDV 19
DB 304 TWPDSNPFVNOGLYGRYIDV 323
XX
XX RESULT 44
XX ABP29749
XX ID ABP29749 standard; Protein: 1252 AA.
XX AC ABP29749;
XX DT 02-JUL-2002 (first entry)
XX DE Streptococcus polypeptide SEQ ID NO 8674.
XX XX
XX KM Streptococcus; GAS; GBS; group B streptococcus; Streptococcus agalactiae;
XX KM group A streptococcus; Streptococcus pyogenes; antibacterial;
XX KM antiinflammatory; infection; vaccine; meningitis; gene therapy.
XX OS Streptococcus agalactiae.
XX XX
XX PN WO200234771-A2.
XX PD 02-MAY-2002.
XX PF 29-OCT-2001; 2001WO-GB04789.
XX XX
XX PR 27-OCT-2000; 2000GB-0026333.
XX PR 24-NOV-2000; 2000GB-0028727.
XX PR 07-MAR-2001; 2001GB-0005640.
XX XX
XX PA (CHIR-) CHIRON SPA.
XX PA (GEMO-) INST GENOMIC RES.
XX PI Telford J, Maignani V, Margarit Ros YI, Grandi G, Fraser C;
XX PI Tettein H;
XX DR WPI: 2002-352536/38.
XX DR N-PSDB: ABN70380.

XX XX New Streptococcus protein for the treatment or prevention of infection
XX PT or disease caused by Streptococcus bacteria, such as meningitis, and
XX PT for detecting a compound that binds to the protein -
XX PS Claim 1; Page 3982; 4525pp; English.
XX XX The invention relates to a protein (ABP25413-ABP30895) from group B
XX streptococcus/GBS (Streptococcus agalactiae) or group A streptococcus/GAS
XX (Streptococcus pyogenes), comprising one of 5483 sequences (S1), given in
XX the specification. The proteins have antibacterial and antiinflammatory
XX activity. (1), nucleic acids encoding (1), ABN66044-ABN71526 and
XX antibodies that bind (1) are used in the manufacture of medicaments for
XX the treatment or prevention of infection or disease caused by
XX Streptococcus bacteria, particularly S. agalactiae and S. pyogenes.
XX Nucleic acids encoding (1) are used to detect Streptococcus in a
XX biological sample. (1) is used to determine whether a compound binds to
XX (1). A composition comprising (1) or a nucleic acid encoding (1), may be
XX used as a vaccine or diagnostic composition. The disease caused by
XX Streptococcus that is prevented or treated may be meningitis. Nucleic
XX acid encoding (1) may be used to recombinantly produce (1) and may be
XX used in gene therapy. Antibodies to (1) are used for affinity
XX chromatography, immunoassays, and distinguishing/identifying
XX Streptococcus proteins.
XX SQ Sequence 1252 AA;
XX
XX Query Match 29.2%; Score 48.5; DB 23; Length 1252;
XX Best Local Similarity 50.0%; Pred. No. 1.7e+02;
XX Matches 10; Conservative 5; Mismatches 4; Indels 1; Gaps 1;
OY 1 SWRSGSSFVQG-LFGQHPDV 19
DB 304 TWPDSNPFVNOGLYGRYIDV 323
XX
XX RESULT 45
XX ABB49146
XX ID ABB49146 standard; Protein: 240 AA.
XX AC ABB49146;
XX DT 05-FEB-2002 (first entry)
XX DE Listeria monocytogenes protein #1850.
XX DE Antibacterial; gene therapy; vaccine; biosynthesis; biodegradation;
XX KM vitamin B12; bacterial infection; disease.
XX XX
XX OS Listeria monocytogenes.
XX PN WO200177335-A2.
XX PD 18-OCT-2001.
XX PF 11-APR-2001; 2001WO-FR01118.
XX PR 11-APR-2000; 2000FR-0004629.
XX XX
XX PA (INSP) INST PASTEUR.
XX PI Buchrieser C, Franquet L, Couve E, Rusniok C, Fsihi H, Dehoux P;
XX PI Dussangot O, Chetouani F, Nedjari H, Glaser P, Kunst F, Cossart P;
XX PI Daniels J, Goebel W, Kieft J, Kuhn M, Ng E, Vazquez-Boland JA;
XX PI Dominguez-Bernal G, Garrido-Garcia P, Tierrez-Martinez A, Amend A;
XX PI Chakraborty T, Domann E, Hain T, Berche P, Charbit A, Durant L;
XX PI Perez-Diaz J, Baquero F, Garcia Del Portillo F, Gomez-Lopez N;
XX PI Madueno E, De Pablos B, Wehland J, Kaerst U, Entian K, Hauf J;
XX PI Rose M, Voss H;
XX DR WPI: 2002-010914/01.
XX DR Genomic sequence for Listeria monocytogenes, useful e.g. for treatment

CC invention is also useful for the identification of potential new targets
 CC for antibiotic development. The antisense nucleic acids can also be used
 CC to identify proteins used in proliferation, to express these proteins,
 CC and to obtain antibodies capable of binding to the expressed proteins.
 CC The proteins can be used to screen compounds in rational drug discovery
 CC programmes. The antisense nucleic acid sequence is also useful to screen
 CC for homologous nucleic acids which are required for cell proliferation in
 CC a wide variety of organisms. The present sequence represents an
 CC essential prokaryotic cellular proliferation protein.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 460 AA:

Query Match 28.9%; Score 48; DB 22; Length 460;
 Best Local Similarity 33.3%; Pred. No. 63;
 Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 5 GSSFWGQLFGQHPDYFYIMEPAWH 28

DB 169 GKEMHEMGMLPFAHIDEPIWY 192

RESULT 48

ID AAU82710 standard; Protein; 753 AA.

AC AAU82710;

DT 23-APR-2002 (first entry)

DE Amino acid sequence of novel human protease #9.

XX Human; protease; cancer; immune-related disorder; cardiovascular disease;
 XX neuronal-associated disease; metabolic disorder; inflammatory disorder;
 XX nervous system disorder; sexual dysfunction; pain; mood disorder;
 XX hypertension; psychotic disorder; neurological disorder; dyskinesia;
 XX viral infection; human immunodeficiency virus; HIV; non-viral infection;
 XX ocular disease; cytostatic; enzyme.

OS Homo sapiens.

XX WO200200860-A2.

PD 03-JAN-2002.

PF 26-JUN-2001; 2001WO-US20171.

PR 26-JUN-2000; 2000US-214047P.

PA (SUGEN) SUGEN INC.

PI Playman G, Whyte D, Sudarsanam S, Manning G, Caenepeel S;

PI Charaydzak G;

DR WPI; 2002-139913/18.

DR N-PSDB; ABR31752.

PT Nucleic acids encoding novel human proteases, useful for useful for
 PT treating diseases and disorders such as cancers, immune-related
 PT diseases and disorders, cardiovascular disease (e.g. restenosis) and
 PT inflammatory disorders -

PS Claim 6; Fig 2D-E; 313pp; English.

CC The present invention relates to the isolation of novel human
 CC proteases, and the nucleic acids encoding them. The sequences of
 CC the invention are useful for treating diseases and disorders such as
 CC cancers (e.g. breast, colon, lung), immune-related diseases and disorders
 CC (e.g. inflammatory diseases and asthma), cardiovascular diseases
 CC (e.g. restenosis and coronary thrombosis), brain or neuronal-associated

CC diseases, metabolic disorders (e.g. diabetes, obesity), inflammatory
 CC disorders (e.g. rheumatoid arthritis and psoriasis), central or
 CC peripheral nervous system diseases, migraines, pain, sexual dysfunction,
 CC mood disorders, attention disorders, cognition disorders, hypotension,
 CC hypertension, psychotic disorders, neurological disorders
 CC (e.g. Alzheimer's disease, Parkinson's disease) and dyskinesias.
 CC The nucleic acids and polypeptides are also useful for treating viral
 CC infections caused by human immunodeficiency virus (HIV), and non-viral
 CC infections such as ocular disease (e.g. glaucoma) and macular
 CC degeneration. AAU82702-AAU82760 represent the novel human proteases of
 CC the invention.

SQ Sequence 753 AA:

Query Match 28.9%; Score 48; DB 23; Length 753;
 Best Local Similarity 70.0%; Pred. No. 11e+02;
 Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 19 VEYIMEPAWH 28

DB 99 VFYLPQPSWH 108

RESULT 49

ID ABB66979 standard; Protein; 817 AA.

AC ABB66979;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 27729.

XX Drosophila; developmental biology; cell signalling; insecticide;
 XX pharmaceutical.

XX Drosophila melanogaster.

OS WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

PI WPI; 2001-656860/75.

DR N-PSDB; ABL11082.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 27729; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 XX capable of detecting 1000 or more genes from Drosophila. The invention is
 XX useful in developmental biology and in elucidating cell signalling and
 XX cell-cell interactions in higher eukaryotes for the development of
 XX insecticides, therapeutics and pharmaceutical drugs. The invention
 XX discloses genomic DNA sequences (AB116176-AB130511), expressed DNA
 XX sequences (AB101840-AB116175) and the encoded proteins

CC (AB157737-AB172072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.

SQ Sequence 817 AA;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:11 / Search time 46.4 Seconds
(without alignments)
60.084 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166

Sequence: 1 SWRSGSFVGQLFGQHPDVFYLMPEAMHV 29

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

1: PIR1: *
2: PIR2: *
3: PIR3: *
4: PIR4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	134	80.7	484	2 UC7350	N-acetylglucosamin
2	134	80.7	486	2 UC7351	N-acetylglucosamin
3	124	74.7	484	2 JE0261	N-acetylglucosamin
4	101	60.8	458	2 A57397	chondroitin 6-sulf
5	57	34.3	1570	2 AC2012	hypothetical prote
6	53	31.9	277	2 G89833	teichoic acid tran
7	53	31.9	390	2 G69044	N-acetylornithine
8	52	31.3	518	2 B83472	hypothetical prote
9	52	31.3	614	2 F97381	hypothetical prote
10	52	31.3	614	2 AE2599	hypothetical prote
11	52	31.3	641	2 AG2164	hypothetical prote
12	50	30.1	1051	2 T48933	hypothetical prote
13	50	30.1	1645	2 A37792	WD repeat domain p
14	49.5	29.8	634	2 C83530	spectrin beta H ch
15	49	29.5	4063	2 T42993	potassium uptake p
16	49	29.5	4101	2 T23630	probable spectrin
17	48.5	29.2	449	2 T23636	hypothetical prote
18	48	28.9	240	2 A11305	hypothetical prote
19	48	28.9	240	2 A11677	metal cations ABC
20	48	28.9	245	2 D70670	hypothetical prote
21	48	28.9	338	2 S73902	UDP-glucose 4-epim
22	48	28.9	443	2 T10801	GDP dissociation i
23	48	28.9	460	2 E83617	probable aminotran
24	48	28.9	478	2 F70322	glutamyl-tRNA (Gln
25	48	28.9	518	2 T23120	hypothetical prote
26	47.5	28.6	495	2 D71307	conserved hypotnet
27	47	28.3	176	2 G73781	hypothetical prote
28	47	28.3	284	2 A83726	phosphatidate cycl
29	47	28.3	288	2 F84293	hypothetical prote

30	47	28.3	296	2 H93338	regulatory protein
31	47	28.3	296	2 AB2944	transcription regu
32	47	28.3	349	2 F86649	dtbp-glucose 4,6-d
33	47	28.3	363	2 T20608	hypothetical prote
34	47	28.3	381	2 T15140	hypothetical prote
35	47	28.3	444	2 D98317	probable pyridoxal
36	47	28.3	444	2 AH2965	hypothetical prote
37	47	28.3	465	2 S76464	hypothetical prote
38	47	28.3	465	2 S51613	hypothetical prote
39	47	28.3	701	2 C83052	cyclin homolog CLN
40	47	28.3	1340	2 D96521	polynucleotide
41	47	28.3	3013	2 AB0480	protein F21D18.16
42	46.5	28.0	251	1 ESDKTM	probable invasin y
43	46.5	28.0	251	1 I50520	oleoyl-lacetyl-carri
44	46.5	28.0	542	2 A28550	S-acyl fatty acid
45	46.5	28.0	634	2 F82623	cyclohexanone mono
46	46.5	28.0	820	2 H86246	potassium uptake p
47	46	27.7	175	2 T08948	hypothetical prote
48	46	27.7	234	2 T46754	Adcc protein limpo
49	46	27.7	234	2 F98118	hypothetical prote
50	46	27.7	234	2 H95253	zinc ABC transport
51	46	27.7	276	2 S57627	avirulence protein
52	46	27.7	291	2 F83504	probable transcrip
53	46	27.7	364	2 C75597	thymidine diphosph
54	46	27.7	426	2 T48424	hypothetical prote
55	46	27.7	557	2 E86850	hypothetical prote
56	46	27.7	764	2 D83870	ATP-dependent RNA
57	46	27.7	1331	2 T18310	receptor-adenylate
58	45.5	27.4	380	2 T19280	hypothetical prote
59	45.5	27.4	423	1 XNCBDM	aspartate transami
60	45.5	27.4	430	1 XNRDTM	aspartate transami
61	45.5	27.4	433	2 S01174	aspartate transami
62	45.5	27.4	519	2 D65178	kup protein - Esch
63	45.5	27.4	622	2 A49345	K(+)-uptake protei
64	45.5	27.4	622	2 AG0952	membrane transport
65	45.5	27.4	809	2 B96314	probable quinate d
66	45.5	27.4	809	2 A12968	hypothetical prote
67	45.5	27.4	844	2 S75547	ethylene response
68	45	27.1	214	2 S77391	hypothetical prote
69	45	27.1	218	2 E87552	conserved hypotnet
70	45	27.1	225	2 G83244	probable two-compo
71	45	27.1	225	2 AP2049	ATP-binding protei
72	45	27.1	313	2 D95866	probable transcrip
73	45	27.1	338	2 E72334	conserved hypotnet
74	45	27.1	362	2 A43316	chorismate synthas
75	45	27.1	370	2 G83219	probable dihydrol
76	45	27.1	421	2 C66806	unknown protein T5
77	45	27.1	465	2 A82438	amino transferase,
78	45	27.1	622	2 S65539	GABA/beta-alanine
79	45	27.1	644	1 F6H0A	fibriinogen alpha c
80	45	27.1	648	2 T35120	hypothetical prote
81	45	27.1	669	2 T51220	hypothetical prote
82	45	27.1	808	2 T15562	hypothetical prote
83	45	27.1	866	2 D44234	hypothetical prote
84	45	27.1	884	2 C86684	fibriinogen alpha c
85	45	27.1	1279	2 A12067	prophage pil prote
86	45	27.1	1654	2 A12067	two-component sens
87	45	27.1	1691	1 D54689	protein-tyrosine-p
88	45	27.1	1894	2 C54689	protein-tyrosine-p
89	45	27.1	1912	2 A56178	protein-tyrosine-p
90	44.5	26.8	2505	1 XYRTFA	enoyl-lacetyl-carri
91	44.5	26.8	401	2 F87196	probable membrane
92	44.5	26.8	430	1 XNHUIM	aspartate transami
93	44.5	26.8	430	2 AB3237	aspartate transami
94	44.5	26.8	469	2 S35960	aspartate transami
95	44.5	26.8	483	2 T01128	hypothetical prote
96	44.5	26.8	483	2 G75392	glycosyl hydrolase
97	44.5	26.8	687	2 T27683	hypothetical prote
98	44.5	26.8	1266	2 G97477	hypothetical prote
99	44.5	26.8	1266	2 A52695	conserved hypotnet
100	44	26.5	1396	1 VCBE40	major capsid prote
101	44	26.5	207	2 AG0753	colanic acid capsu
102	44	26.5	218	2 E84840	hypothetical prote
	44	26.5	254	2 S75676	probable phosphoes

249	42	25.3	1816	2	F83901	hypothetical prote	322	41	24.7	416	2	T02194	probable pectinase
250	41.5	25.0	175	2	F86439	hypothetical prote	323	41	24.7	416	2	B95254	d1d2 protein limpo
251	41.5	25.0	184	2	A86461	hypothetical prote	324	41	24.7	427	2	B95118	hypothetical prote
252	41.5	25.0	189	1	LEVCTP	thermostable direc	325	41	24.7	439	2	E97079	sugar-binding peri
253	41.5	25.0	250	2	T40377	cdp-diacylglycerol	326	41	24.7	441	2	B84854	hypothetical prote
254	41.5	25.0	304	2	B91034	probable fructokin	327	41	24.7	446	2	G70774	probable atph prot
255	41.5	25.0	305	1	S52161	probable fructokin	328	41	24.7	452	2	T28094	hypothetical prote
256	41.5	25.0	305	2	C83878	D-fructokinase [lm	329	41	24.7	461	2	AH2923	hypothetical prote
257	41.5	25.0	346	2	G75069	ADP-heptose-LPS he	330	41	24.7	463	2	T45074	Sun protein limpor
258	41.5	25.0	346	2	F82349	ADP-heptose-LPS he	331	41	24.7	463	2	H97697	probable phenylala
259	41.5	25.0	348	2	S34494	ADP-heptose-LPS he	332	41	24.7	472	2	C42391	probable sun prote
260	41.5	25.0	410	1	IXBE10	alpha trans-induci	333	41	24.7	476	2	E83006	Ca2+-transporting
261	41.5	25.0	437	2	S15306	CDP-4-keto-6-deoxy	334	41	24.7	477	2	T33531	two-component resp
262	41.5	25.0	482	2	C82404	6-phosphogluconate	335	41	24.7	495	2	T05388	hypothetical prote
263	41.5	25.0	485	2	S73655	anthranilate synth	336	41	24.7	496	2	T49698	hypothetical prote
264	41.5	25.0	504	2	A82205	probable deoxyribo	337	41	24.7	502	2	C84400	hypothetical prote
265	41.5	25.0	510	2	JC4208	nitrogenase NifB c	338	41	24.7	514	2	AC3563	phenylalanyl-tRNA
266	41.5	25.0	537	2	T09237	nifb protein - Fra	339	41	24.7	530	2	A90741	ribose transport A
267	41.5	25.0	537	2	S78195	C4-dicarboxylate t	340	41	24.7	530	2	D85591	probable ABC-type
268	41.5	25.0	597	2	B82140	cytochrome-c oxida	341	41	24.7	530	2	D64819	probable ABC-type
269	41.5	25.0	630	2	F83876	cytochrome-c oxida	342	41	24.7	531	2	AD0602	probable ABC-type
270	41.5	25.0	622	2	E86609	cytochrome-c oxida	343	41	24.7	536	2	S40717	hypothetical prote
271	41.5	25.0	635	2	D45335	cytochrome-c oxida	344	41	24.7	540	2	B45665	adult-specific ol.
272	41.5	25.0	642	2	T51421	L-aspartate oxida	345	41	24.7	550	2	A48053	hypothetical prote
273	41.5	25.0	706	2	S62501	hypothetical prote	346	41	24.7	553	2	F86202	ABC transporter AT
274	41.5	25.0	764	2	D84847	probable receptor-	347	41	24.7	589	2	A81893	phenylalanyl-tRNA
275	41.5	25.0	802	2	G87578	TonB-dependent rec	348	41	24.7	622	2	A40144	hypothetical prote
276	41.5	25.0	823	2	D81129	calton transport A	349	41	24.7	636	2	C81128	prolactin receptor
277	41.5	25.0	944	2	T47246	chitin synthase (E	350	41	24.7	636	2	B81128	probable ABC trans
278	41.5	25.0	993	2	C82687	glycine decarboxyl	351	41	24.7	641	2	A24075	ABC transporter, A
279	41.5	25.0	1215	2	T32734	myosin-1A - Acanth	352	41	24.7	645	2	G01205	lipase precursor -
280	41.5	25.0	1400	2	T23644	hypothetical prote	353	41	24.7	669	2	E71127	tyl protein - huma
281	41.5	25.0	2396	2	AC0304	probable hemolysin	354	41	24.7	671	2	H86702	hypothetical prote
282	41.5	25.0	2535	2	T13714	hypothetical prote	355	41	24.7	843	1	A27340	potassium uptake p
283	41.5	25.0	2833	2	T23064	protein T22A3.8 [l	356	41	24.7	864	1	JC1422	complement C7 prec
284	41.5	25.0	2823	2	F87908	laminin alpha chai	357	41	24.7	884	2	T40690	nitrate reductase
285	41.5	25.0	3102	2	T43291	hypothetical prote	358	41	24.7	891	2	B48642	hypothetical prote
286	41.5	25.0	3152	2	G71955	probable acetyltra	359	41	24.7	925	2	T01384	acornulate hydratas
287	41.5	25.0	3177	2	T37137	plasma membrane ca	360	41	24.7	934	2	161714	hypothetical prote
288	41.5	25.0	3178	2	G03294	hypothetical prote	361	41	24.7	990	2	T03784	co-repressor prote
289	41.5	25.0	379	2	S75860	hypothetical prote	362	41	24.7	1016	1	J50428	probable receptor
290	41.5	25.0	183	2	F69049	aequorin precursor	363	41	24.7	1060	2	F88710	NAD ADP-ribosyltra
291	41.5	25.0	196	1	AQJPNV	aequorin precursor	364	41	24.7	1079	2	T30956	protein C01G5.4 [l
292	41.5	25.0	196	1	AQJPNV	aequorin precursor	365	41	24.7	1403	2	H67131	hypothetical prote
293	41.5	25.0	205	2	AD0774	probable exported	366	41	24.7	1449	2	B31963	uncharacterized, p
294	41.5	25.0	224	2	T01185	protein kinase (EC	367	41	24.7	2064	2	T13707	IgA-specific serin
295	41.5	25.0	228	2	C81243	conserved hypotet	368	41	24.7	2242	2	A57541	still life protein
296	41.5	25.0	230	2	JC1483	deoxyribonuclease	369	41	24.7	108	2	S21420	pyrimidine synthe
297	41.5	25.0	234	2	H89920	conserved hypotet	370	41	24.4	156	2	S67248	general stress pro
298	41.5	25.0	241	2	H81278	phosphatidate cyti	371	41	24.4	185	2	JC7369	hypothetical prote
299	41.5	25.0	247	1	D24706	nodulation protein	372	41	24.4	242	2	A10659	phenylated Rab acc
300	41.5	25.0	247	1	H95319	phosphadenyl-yl-su	373	41	24.4	252	2	A12088	conserved hypotet
301	41.5	25.0	254	2	D84559	probable peptide m	374	41	24.4	266	2	A5186	acyl-lacyl-carrier
302	41.5	25.0	279	2	F64109	diethylsulfoxide	375	41	24.4	269	2	T20670	acyl-lacyl-carrier
303	41.5	25.0	280	2	G81651	prolactin receptor	376	41	24.4	270	2	H64433	hypothetical prote
304	41.5	25.0	288	2	B59405	prolactin receptor	377	41	24.4	280	2	E82059	hypothetical prote
305	41.5	25.0	301	2	H82569	probable GDSL-moli	378	41	24.4	286	2	T21545	phosphatidate cyti
306	41.5	25.0	303	2	F84860	hypothetical prote	379	41	24.4	301	2	G83352	hypothetical prote
307	41.5	25.0	312	2	S51440	hypothetical prote	380	41	24.4	322	2	T50557	transcription regu
308	41.5	25.0	314	2	T38955	hypothetical prote	381	41	24.4	332	2	S49034	lactate dehydrogen
309	41.5	25.0	317	2	F85650	hypothetical prote	382	41	24.4	333	2	A47160	tripeptide dehydrog
310	41.5	25.0	330	2	C95844	hypothetical prote	383	41	24.4	335	2	G82247	cysteine synthase/
311	41.5	25.0	330	2	D87068	probable sugar ABC	384	41	24.4	336	2	G84404	hypothetical prote
312	41.5	25.0	340	2	A82218	transcription regu	385	41	24.4	376	2	E83409	hypothetical prote
313	41.5	25.0	349	2	T01417	transcription regu	386	41	24.4	389	2	AE2293	hypothetical prote
314	41.5	25.0	356	2	H70882	hypothetical prote	387	41	24.4	399	2	A84475	hypothetical prote
315	41.5	25.0	369	2	A59405	maltoase/maltodextr	388	41	24.4	435	2	A84824	hypothetical prote
316	41.5	25.0	376	2	A59405	maltoase/maltodextr	389	41	24.4	435	2	A84824	hypothetical prote
317	41.5	25.0	383	2	C69442	prolactin receptor	390	41	24.4	457	2	T04444	cytochrome p450 -
318	41.5	25.0	389	2	A87602	succinyl-CoA synth	391	41	24.4	485	2	C70488	cytochrome-c oxida
319	41.5	25.0	400	2	AF0291	conserved hypotet	392	41	24.4	496	2	G89824	hypothetical prote
320	41.5	25.0	401	2	T39185	probable drug resi	393	41	24.4	499	2	A38891	cytochrome-c oxida
321	41.5	25.0	411	2	T15705	probable polysacch	394	41	24.4	512	2	S28663	cytochrome-c oxida
						hypothetical prote				528	2	JC4814	Na+/H+-exchanging

395	40.5	24.4	540	2	E69861	ABC transporter (A	468	40	24.1	448	2	C97757	hypothetical prote
396	40.5	24.4	549	2	A87347	hypothetical prote	469	40	24.1	454	1	A26955	alkaline serine pr
397	40.5	24.4	552	2	AC3435	cytochrome-c oxida	470	40	24.1	461	1	DCMSO	ornithine decarbox
398	40.5	24.4	556	2	C87472	steroid monooxygen	471	40	24.1	461	2	I55356	ornithine decarbox
399	40.5	24.4	600	2	T48433	hypothetical prote	472	40	24.1	461	2	I56477	ornithine decarbox
400	40.5	24.4	616	2	JX0140	cytochrome-c oxida	473	40	24.1	462	2	T30164	probable phosphor
401	40.5	24.4	642	2	A11827	cyclomaltodextrin	474	40	24.1	467	2	AE3142	hypothetical prote
402	40.5	24.4	673	2	T47006	hypothetical prote	475	40	24.1	468	2	A69611	cytochrome bd ubiq
403	40.5	24.4	673	2	AC0238	heamin storage sys	476	40	24.1	468	2	AD1790	cytochrome D ubiq
404	40.5	24.4	729	2	F86308	similar to disease	477	40	24.1	468	2	AE1414	125k surface anti
405	40.5	24.4	729	2	T40080	probable ABC trans	478	40	24.1	471	2	PS0154	phosphomannomutase
406	40.5	24.4	749	2	T43370	oligosaccharyltran	479	40	24.1	473	2	S22621	D-alanyl-D-alanine
407	40.5	24.4	752	2	T39338	oligosaccharyl tra	480	40	24.1	477	2	E91136	D-alanyl-D-alanine
408	40.5	24.4	752	2	S56146	GCN20 protein - ye	481	40	24.1	477	2	H85981	serine-type D-Ala-
409	40.5	24.4	763	2	A47563	glucose-6-phosphat	482	40	24.1	477	2	AE4535	hypothetical prote
410	40.5	24.4	831	2	S62751	probable DNA-direc	483	40	24.1	484	2	E84765	probable apcC prot
411	40.5	24.4	878	1	T01774	hypothetical prote	484	40	24.1	485	2	E70558	regulatory protein
412	40.5	24.4	1180	2	NCEKX5	exodeoxyribonuclea	485	40	24.1	495	2	S60175	probable secreted
413	40.5	24.4	1180	2	E91088	DNA helicase RecB	486	40	24.1	497	2	AF0505	probable cytochrom
414	40.5	24.4	1180	2	G85933	hypothetical prote	487	40	24.1	501	2	T36051	conserved hypotet
415	40.5	24.4	1180	2	E82754	hypothetical prote	488	40	24.1	509	2	D82212	probable aminotran
416	40.5	24.1	88	2	E95414	hypothetical prote	489	40	24.1	520	2	G98145	cytochrome bd oxid
417	40.5	24.1	122	2	S57090	hypothetical prote	490	40	24.1	520	2	C81424	hypothetical prote
418	40.5	24.1	166	2	A75426	hypothetical prote	491	40	24.1	529	2	S76831	hypothetical prote
419	40.5	24.1	189	2	T23747	transposon-related	492	40	24.1	532	2	AC2112	probable homeobox
420	40.5	24.1	200	2	B67700	phosphoglycerate m	493	40	24.1	532	2	T05281	ABC transporter, p
421	40.5	24.1	212	1	AB2419	coat protein - pap	494	40	24.1	535	2	F90418	mark protein (trk
422	40.5	24.1	215	1	VCWGPB	hypothetical prote	495	40	24.1	544	2	S58532	phosphoglucosylase
423	40.5	24.1	222	2	B84410	hypothetical prote	496	40	24.1	545	2	D87259	probable protein k
424	40.5	24.1	222	2	JC1199	alpha-glucosidase	497	40	24.1	546	2	D84555	lysyl-tRNA synthet
425	40.5	24.1	235	2	D65928	probable holoctoc	498	40	24.1	552	2	C87259	hypothetical prote
426	40.5	24.1	243	2	G95928	hypothetical prote	499	40	24.1	554	2	AH2321	maltese ABC transp
427	40.5	24.1	245	2	F91253	probable DNA motif	500	40	24.1	564	2	A72283	ABC transporter AT
428	40.5	24.1	249	1	SYECDG	phosphatidate cyti	501	40	24.1	566	2	D84872	GTPase-activating
429	40.5	24.1	249	1	A95651	CDP-diglyceride sy	502	40	24.1	615	2	B39697	olipeptide ABC t
430	40.5	24.1	253	2	AG5502	hypothetical prote	503	40	24.1	633	2	D72284	hypothetical prote
431	40.5	24.1	253	2	AG5502	hypothetical prote	504	40	24.1	651	2	AF2493	GTPase-activating
432	40.5	24.1	263	1	A26525	oleoyl-lacetyl-carri	505	40	24.1	663	2	A39897	conserved hypotet
433	40.5	24.1	268	2	E75607	2-oxo-hepta-3-ene-	506	40	24.1	664	2	F83376	cyclomaltodextrin
434	40.5	24.1	272	2	G70324	prolipo-protein dia	507	40	24.1	718	1	ALBSMX	hypothetical prote
435	40.5	24.1	278	2	B83453	hypothetical prote	508	40	24.1	718	2	H72519	glycosyltransferas
436	40.5	24.1	282	2	A10128	phosphatidate cyti	509	40	24.1	732	2	A84107	hypothetical prote
437	40.5	24.1	285	2	AH0529	pyridoxal kinase (510	40	24.1	733	2	S76856	hypothetical prote
438	40.5	24.1	286	2	A10288	conserved hypotet	511	40	24.1	733	2	D84854	UvrA-like ABC tran
439	40.5	24.1	296	2	D87525	dehydratogenase/redu	512	40	24.1	752	2	T35244	secreted protease
440	40.5	24.1	299	2	H83749	hypothetical prote	513	40	24.1	781	2	G96991	NADH dehydrogenas
441	40.5	24.1	322	2	AG1929	hypothetical prote	514	40	24.1	787	2	H70374	coat protein - lei
442	40.5	24.1	329	2	D83778	hypothetical prote	515	40	24.1	791	2	B46171	hypothetical prote
443	40.5	24.1	330	2	T45981	hypothetical prote	516	40	24.1	794	2	S50687	probable dehydroge
444	40.5	24.1	336	2	G71151	probable dTPP-giuc	517	40	24.1	825	2	F95963	hypothetical prote
445	40.5	24.1	339	2	T34406	hypothetical prote	518	40	24.1	829	2	B96640	hypothetical prote
446	40.5	24.1	343	2	E69788	DNA restriction ho	519	40	24.1	840	2	T21333	pol polyprotein -
447	40.5	24.1	343	2	E69678	involved in polyke	520	40	24.1	843	2	S33123	multifunctional be
448	40.5	24.1	346	2	T08864	hypothetical prote	521	40	24.1	894	2	S54786	genome polyprotein
449	40.5	24.1	348	2	G02297	gene N33 protein -	522	40	24.1	929	2	A44048	chordin precursor
450	40.5	24.1	359	2	T16350	hypothetical prote	523	40	24.1	941	1	A53195	hypothetical WD-40
451	40.5	24.1	370	2	T35416	hypothetical prote	524	40	24.1	962	2	S58107	env polyprotein -
452	40.5	24.1	371	2	T40287	hypothetical prote	525	40	24.1	985	1	VCLJSE	light junction pro
453	40.5	24.1	388	2	C82701	lipic acid synthe	526	40	24.1	994	2	S18739	probable membrane
454	40.5	24.1	388	2	A82903	conserved hypotet	527	40	24.1	1163	2	JEO366	major capsid prote
455	40.5	24.1	400	1	ZBBE14	44.1K zinc-binding	528	40	24.1	1246	1	VCBE17	protoporphyrin IX
456	40.5	24.1	400	1	H87444	hypothetical prote	529	40	24.1	1374	2	T01789	hypothetical prote
457	40.5	24.1	403	2	T08471	harpin - Erwina a	530	40	24.1	1382	2	A110627	glutamate synthase
458	40.5	24.1	404	2	AE4211	aspartate transami	531	40	24.1	1396	2	A89813	spectrin beta-6 ch
459	40.5	24.1	407	2	JC5124	hypothetical prote	532	40	24.1	1439	2	A44159	utrophin - human
460	40.5	24.1	418	2	C72394	probable lipo prot	533	40	24.1	1463	1	S28381	tegument protein 2
461	40.5	24.1	420	2	E70914	probable PPE prote	534	40	24.1	1463	1	T42567	hypothetical prote
462	40.5	24.1	423	2	C70582	probable integral	535	40	24.1	1463	2	AH2515	hypothetical prote
463	40.5	24.1	427	2	D81784	AMG-related prote	536	40	24.1	1463	2	B95418	hypothetical prote
464	40.5	24.1	428	2	B69187	conserved hypotet	537	40	24.1	1463	2	B83756	stellate protein -
465	40.5	24.1	432	2	A90465	hypothetical prote	538	40	24.1	1463	2	S24397	hypothetical prote
466	40.5	24.1	432	2	B70776	probable gln2 - My	539	40	24.1	1463	2	B72298	hypothetical prote
467	40.5	24.1	446	2			540	40	24.1	1463	2		

979	38	22.9	392	2	B6549	polymorphic outer
980	38	22.9	392	2	B12075	polymorphic outer
981	38	22.9	394	2	T45672	hypothetical prote
982	38	22.9	395	2	A12412	hypothetical prote
983	38	22.9	396	2	T50229	probable transmem
984	38	22.9	397	2	S28274	hypothetical prote
985	38	22.9	398	2	D66700	hypothetical prote
986	38	22.9	399	2	F70937	hypothetical prote
987	38	22.9	400	2	A26258	endoplasmic retic
988	38	22.9	401	1	S10528	5-aminolevulinat
989	38	22.9	401	1	T21288	phosphoprotein pho
990	38	22.9	401	1	A30696	probable integral
991	38	22.9	402	2	T14710	probable transpos
992	38	22.9	402	2	A30193	transposase, IS285
993	38	22.9	402	2	A30356	transposase, IS285
994	38	22.9	402	2	A30472	transposase, IS285
995	38	22.9	402	2	A30341	transposase, IS285
996	38	22.9	402	2	A30002	transposase, IS285
997	38	22.9	402	2	A30190	transposase, IS285
998	38	22.9	402	2	A30242	transposase, IS285
999	38	22.9	402	2	A30267	transposase, IS285
1000	38	22.9	402	2	A30348	transposase, IS285

ALIGNMENTS

RESULT 1

UC7350 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
 C:Accession: J07350
 R:Uchimura, K.; Fasekhan, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa, N.;
 Biochem. Biophys. Res. Commun. 274, 291-296, 2000
 A>Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
 A:Reference number: UC7350; MUID:20374462; PMID:10913333
 A:Accession: J07350
 A:Molecule type: mRNA
 A:Residues: 1-484 <UCH>
 A:Cross-references: DDBJ:AB040710
 C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat
 sulfate linkage formation, this enzyme is involved in synthesis of L-selectin ligand in
 C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match Best Local Similarity 72.4%; Score 134; DB 2; Length 484;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAWY 29

DB 107 TWRGSSFLGELFNOHPDVFYIMEPAWY 135

RESULT 2

J07351 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
 C:Species: Homo sapiens (man)
 C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
 C:Accession: J07351
 R:Uchimura, K.; Fasekhan, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa, N.;
 Biochem. Biophys. Res. Commun. 274, 291-296, 2000
 A>Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of a
 A:Reference number: J07350; MUID:20374462; PMID:10913333
 A:Accession: J07351
 A:Molecule type: mRNA
 A:Residues: 1-486 <UCH>
 A:Cross-references: DDBJ:AB040711
 C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosylat
 sulfate linkage formation, this enzyme is involved in synthesis of L-selectin ligand in
 C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 80.7%; Score 134; DB 2; Length 486;

Best Local Similarity 72.4%; Pred. No. 3.2e-11;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAWY 29

DB 109 TWRGSSFLGELFNOHPDVFYIMEPAWY 137

RESULT 3

JE0261 N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
 C:Species: Homo sapiens (man)
 C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
 C:Accession: JE0261
 R:Uchimura, K.; Muramatsu, H.; Kaneko, T.; Ogawa, H.; Yamakawa, T.; Fan, Q.W.; Mitsuo
 T.;
 J. Biochem. 124, 670-678, 1998
 A>Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis
 A:Reference number: JE0261; MUID:98391845; PMID:9722682
 A:Accession: JE0261
 A:Molecule type: mRNA
 A:Residues: 1-484 <UCH>
 A:Cross-references: DDBJ:AB014679
 C:Comment: This protein catalyzes the transfer of sulfur of sulfate from 3'-phosphoad
 C:Superfamily: chondroitin 6-sulfotransferase
 C:Keywords: sulfotransferase

Query Match Best Local Similarity 74.7%; Score 124; DB 2; Length 484;
 Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SMRGSFVGOLFQHPDVFYIMEPAWY 29

DB 126 TWRGSSFLGELFNOHPDVFYIMEPAWY 154

RESULT 4

A57397 chondroitin 6-sulfotransferase (EC 2.8.2.17) precursor - chicken
 C:Species: Gallus gallus (chicken)
 C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 20-Jun-2000
 C:Accession: A57397
 R:Uchimura, K.; Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuc
 J. Biol. Chem. 270, 18575-18580, 1995
 A>Title: Molecular cloning and expression of chick chondrocyte chondroitin 6-sulfot
 A:Reference number: A57397; MUID:95355450; PMID:7629189
 A:Accession: A57397
 A>Status: Preliminary
 A:Molecule type: mRNA
 A:Residues: 1-458 <EUC>
 A:Cross-references: GB:D49915; NID:9971262; PIDN:BA008655.1; PID:9971263
 C:Superfamily: chondroitin 6-sulfotransferase
 C:Keywords: sulfotransferase

Query Match Best Local Similarity 60.8%; Score 101; DB 2; Length 458;
 Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RCGSFVGFQGFQHPDVFYIMEPAWY 29

DB 122 RTGSSFLGELFNOHPDVFYIMEPAWY 148

RESULT 5

AC2012 hypothetical protein a11649 [imported] - Nostoc sp. (strain PCC 7120)
 C:Species: Nostoc sp.
 A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C>Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
 C:Accession: AC2012
 R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Irigu
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata

AE2599
hypothetical protein Atu0187 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C:Species: Agrobacterium tumefaciens
C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 01-Feb-2002
C:Accession: AE2599
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McCell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; PMID:11743193
A:Accession: AE2599
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-614 <KUR>
A:Cross-references: GB:AE008688; PIDN:AA141211.1; PID:917738513; GSPDB:GN00186
C:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: Atu0187
A:Map position: circular chromosome
C:Superfamily: periplasmic oligopeptide-binding protein

Query Match 31.3%; Score 52; DB 2; Length 614;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

OY 2 WRSGSFVGOLFQHPDV 20
DB 33 WRKGISTVCEL--KHPDGF 49

RESULT 11
AG2164
hypothetical protein all2870 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp.
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AG2164
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasa moto, S.; Watanabe, A.; Iriyuchi
Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Ana
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AG2164
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-641 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA074569.1; PID:917131964; GSPDB:GN00179
C:Genetics:
A:Experimental source: strain PCC 7120
A:Gene: all2870

Query Match 31.3%; Score 52; DB 2; Length 641;
Best Local Similarity 46.7%; Pred. No. 21;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 WRSGSFVGOLFQHPDV 16
DB 298 WRGSGYINSFPGYH 312

RESULT 12
T48933
WD repeat domain protein - Arabidopsis thaliana
N:Alternate names: protein F14L2.80
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C:Accession: T48933
R:Jordan, N.; Bangert, S.; Wiedelmann, R.; Voss, H.; Unsel, M.; Mewes, H.W.; Rudd, S.;
submitted to the Protein Sequence Database, April 2000
A:Reference number: Z25008

A:Accession: T48933
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1051 <OR>
A:Cross-references: EMBL:AL353818; GSPDB:GN00061; ATSP:F14L2.80
A:Experimental source: cultivar Columbia; BAC clone F14L2
C:Genetics:
A:Gene: ATSP:F14L2.80
A:Map position: 3
A:Introns: 36/3; 121/3; 205/3; 336/2; 668/3; 739/1; 773/1; 853/2; 894/2; 932/3

Query Match 30.1%; Score 50; DB 2; Length 1051;
Best Local Similarity 36.7%; Pred. No. 72;
Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

OY 1 WRSGSFVGOLFQHPDV--FYLMEPRMH 28
DB 823 SWNLGAPGELAGLDVDRYKMARKPEWN 852

RESULT 13
A37792
Spectrin beta-H chain - fruit fly (Drosophila melanogaster) (fragment)
C:Species: Drosophila melanogaster
C:Date: 30-Apr-1991 #sequence_revision 08-Nov-1996 #text_change 16-Jul-1999
C:Accession: A37792; S70848; S15666
R:Dubreuil, R.R.; Byers, T.J.; Stewart, C.T.; Kiehart, D.P.
J. Cell Biol. 111, 1849-1858, 1990
A:Title: A beta-spectrin isoform from Drosophila (beta-H) is similar in size to verte
A:Reference number: A37792; MUID:91055599; PMID:2229176
A:Accession: A37792
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-1645 <DUB>
A:Cross-references: GB:X53992
A:Note: Met-14 is the probable initiator
R:Dubreuil, R.R.
submitted to the EMBL Data Library, July 1990
A:Reference number: S70848
A:Accession: S70848
A:Molecule type: mRNA
A:Residues: 1-400/403-410, OL/411-1645 <DUB>
A:Cross-references: EMBL:X53992; MUID:97654; PID:97655
A:Note: Met-14 is the probable initiator
C:Genetics:
A:Gene: FlyBase:kst
A:Cross-references: FlyBase:FBgn0004167
C:Superfamily: alpha-actinin actin-binding domain homology; SH3 homology; spectrin/dy
C:Keywords: actin binding
F:36-258/Domain: alpha-actinin actin-binding domain homology <ACT>
F:288-399/Domain: spectrin/dystrophin repeat homology <SP1>
F:400-510/Domain: spectrin/dystrophin repeat homology <SP2>
F:851-896/Domain: SH3 homology <SH3>
F:990-1091/Domain: spectrin/dystrophin repeat homology <SP3>
F:1303-1408/Domain: spectrin/dystrophin repeat homology <SP4>
F:1409-1512/Domain: spectrin/dystrophin repeat homology <SP5>

Query Match 30.1%; Score 50; DB 2; Length 1645;
Best Local Similarity 45.0%; Pred. No. 1,2e+02;
Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 1 WRSGSFVGOLFQHPDV 20
DB 186 SWRGLGFMNLHSHRPDLF 205

RESULT 14
C83330
potassium uptake protein KUP PA0917 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: C83330
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

QY 8 FVQGLFGQHPDYFYLMKP 25
| : : : : :
DB 143 FTRALAQHAEIFFLDEP 160

RESULT 19

A:Accession: A11677
metal cations ABC transporter, ATP-binding protein homolog 11n1963 [imported] - *Listeria*
C:Species: *Listeria innocua*
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 17-May-2002
C:Accession: A11677
R:Jaaser, P.; Franzen, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A:Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapat, G.; Madueno, E.; Maitournam, A.; Meek, C.; Schlueter, T.; Simoes, N.; Tietzer, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.
Title: Comparative genomics of *Listeria* species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: A11677
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-240 <GLA>
A:Cross-references: GB:AL592022; PIDN:CA097193.1; PID:q16414464; GSPDB:GN00178
A:Experimental source: strain Clp11262
C:Genetics:
A:Gene: 11n1963
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 28.9%; Score 48; DB 2; Length 240;
Best Local Similarity 38.9%; Pred. No. 27;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 8 FVQGLFGQHPDYFYLMKP 25
| : : : : :
DB 143 FTRALAQHAEIFFLDEP 160

RESULT 20

D70670
hypothetical protein RV2959c - *Mycobacterium tuberculosis* (strain H37RV)
N:Alternate names: u0002j2b protein
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70670; S73062
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Kellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70670
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-245 <COL>
A:Cross-references: GB:283018; GB:AL123456; NID:q3261671; PIDN:CA05417.1; PID:e283380;
A:Experimental source: strain H37RV
R:Smith, D.R.; Robison, K.
Submitted to the EMBL Data Library, September 1994
A:Description: *Mycobacterium tuberculosis* cosmid tb02.
A:Reference number: S73053
A:Accession: S73062
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-245 <SMI>
A:Cross-references: EMBL:U00024; NID:q560506; PIDN:AAA50941.1; PID:q560520
C:Genetics:
A:Gene: RV2959c

Query Match 28.9%; Score 48; DB 2; Length 245;
Best Local Similarity 69.2%; Pred. No. 27;
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 2 WRSQSFYQGLNG 14
| : : : : :
DB 5 WRSRTSLVQGLIG 17

RESULT 21

S73902
UDP-glucose 4-epimerase gale - *Mycoplasma pneumoniae* (strain ATCC 29342)
N:Alternate names: hypothetical protein A65_corf38
C:Species: *Mycoplasma pneumoniae*
A:Valley: ATCC 29342
C:Date: 27-Feb-1997 #sequence_revision 25-Apr-1997 #text_change 07-Dec-1999
C:Accession: S73902
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A:Title: Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*
A:Reference number: S73327; MUID:97105885; PMID:8948653
A:Accession: S73902
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-338 <HIM>
A:Cross-references: EMBL:AE000056; GB:U00089; NID:q1674263; PIDN:AA096224.1; PID:q167
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C:Genetics:
A:Gene: gale
A:Genetic code: SGC3
C:Superfamily: *Escherichia coli* UDPglucose 4-epimerase; UDPglucose 4-epimerase homolo
F:8-338/Domain: UDPglucose 4-epimerase homology <UDP>

Query Match 28.9%; Score 48; DB 2; Length 338;
Best Local Similarity 60.0%; Pred. No. 39;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 5 GSSEFYQGLFGQHPDY 19
| : : : : :
DB 18 GSCFIDQLKQYPDY 32

RESULT 22

T10801
GDP dissociation inhibitor GDII - *Volvox carter* f. *nagariensis*
C:Species: *Volvox carter* f. *nagariensis*
C:Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 26-May-2000
C:Accession: T10801
R:Beysse, K.; Fabry, S.
Submitted to the EMBL Data Library, July 1996
A:Description: Identification and characterization of a lower plant Ypt/Rab guanosine
A:Reference number: Z17156
A:Accession: T10801
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-443 <BEY>
A:Cross-references: EMBL:U62866; NID:q1572517; PID:q1572518
A:Experimental source: strain HK10
C:Genetics:
A:Gene: GDII
C:Function:
A:Description: inhibits dissociation of GDP from GMP binding proteins
C:Superfamily: human GDP dissociation inhibitor XAP-4

Query Match 28.9%; Score 48; DB 2; Length 443;
Best Local Similarity 47.4%; Pred. No. 53;
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 8 FVQGLFGQHPDYFYLMKP 26
| : : : : :
DB 183 FIGHALGLHDDAYLTQPA 201

RESULT 23
E83617
probable aminotransferase PA0221 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: E83617
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Badian, S.; Vran, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: E83617
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-460 <STO>
 A:Cross-references: GB:AE004460; GB:AE004091; NID:g9946055; PIDN:AA03610.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA0221
 C:Superfamily: ornithine-oxo-acid aminotransferase

Query Match 28.9%; Score 48; DB 2; Length 460;
 Best Local Similarity 33.3%; Pred. No. 55;
 Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

OY 5 GSSFGQLFGQHPDVFYLMPEAW 28
 DB 169 GAKFMHMGMLPFAHIDEPYMY 192

RESULT 24
 F70322
 gltamyl-tRNA (Gln) amidotransferase subunit A - *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 22-Jun-1999
 C:Accession: F70322
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: F70322
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-478 <AOE>
 A:Cross-references: GB:AE000680; NID:g2982948; PIDN:AA006569.1; PID:g2982954; GB:AE00065
 A:Experimental source: strain VFS
 C:Genetics:
 A:Gene: g8tA
 C:Superfamily: indoleacetamide hydrolase

Query Match 28.9%; Score 48; DB 2; Length 478;
 Best Local Similarity 36.7%; Pred. No. 58;
 Matches 11; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

OY 1 SMRSGSFVGLFGQHPDVFYLMPEAW 24
 DB 437 AKMDGLPVGGQLGKRWDETTLLQISTYLMW 466

RESULT 25
 T23120
 hypothetical protein H25K10.6 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 20-Jun-2000
 R:Morimoto, B.
 A:Title: The EMBL Data Library, March 1997
 A:Reference number: Z19684
 A:Accession: T23120
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-518 <MIL>
 A:Cross-references: EMBL:Z92796; PIDN:CAB07233.1; GSPDB:GN00022; CESP:H25K10.6
 A:Experimental source: clone H25K10

C:Genetics:
 A:Gene: CESP:H25K10.6
 A:Map position: 4
 A:Introns: 59/1; 144/1; 211/1; 307/3; 453/3
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein T05A6.4

Query Match 28.9%; Score 48; DB 2; Length 518;
 Best Local Similarity 26.9%; Pred. No. 63;
 Matches 7; Conservative 8; Mismatches 11; Indels 0; Gaps 0;

OY 2 WRSFGVGLFGQHPDVFYLMPEAW 27
 DB 484 YTTNAFIQDYRGEPYIFFLKCELM 509

RESULT 26
 D71307
 conserved hypothetical integral membrane protein TP0582 - *Syphilis spirochete*
 C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)
 C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 05-Nov-1999
 C:Accession: D71307
 R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, R.; G.
 rison, J.; Kralak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterlidge, T.; M.
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 A:Title: Complete genome sequence of *Treponema pallidum*, the *Syphilis spirochete*.
 A:Reference number: A71250; MUID:98332770; PMID:9665876
 A:Accession: D71307
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-495 <COI>
 A:Cross-references: GB:AE001233; GB:AE000520; NID:g3322870; PIDN:AA065556.1; PID:g332
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0582

Query Match 28.6%; Score 47.5; DB 2; Length 495;
 Best Local Similarity 35.3%; Pred. No. 70;
 Matches 12; Conservative 7; Mismatches 10; Indels 5; Gaps 2;

OY 1 SMRSGSFVGLFGQHPDVFYLMPEAW 29
 DB 416 SMRTAIFAGQLFRTSAHVPTPIHLDPAHYL 449

RESULT 27
 G72781
 hypothetical protein APE0240 - *Aeropyrum pernix* (strain K1)
 C:Species: *Aeropyrum pernix*
 C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
 C:Accession: G72781
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Halkawa, Y.; Jin-no, K.; Ta.
 awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
 A:Title: Complete genome sequence of an aerobic hyperthermophilic *Crenarchaeon*, *Aero*
 A:Reference number: A72450; MUID:9910339; PMID:10382966
 A:Accession: G72781
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-176 <KAN>
 A:Cross-references: DBJ:AP000058; NID:g5103388; PIDN:BA079153.1; PID:g5103632
 A:Experimental source: strain K1
 C:Genetics:
 A:Gene: APE0240
 C:Superfamily: *Aeropyrum pernix* hypothetical protein APE0240

Query Match 28.3%; Score 47; DB 2; Length 176;
 Best Local Similarity 50.0%; Pred. No. 26;
 Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

OY 5 GSSFGVGLFGQHPDVFYLMPEAW 24
 DB 45 GSVRAVSAYFSQHPVEYVNE 64

A.Variety: FCC 6803
C.Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C.Accession: S76464
R.Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;
O, K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Matanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A.Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*
s.
A.Reference number: S74322; MUID:97061201; PMID:8905231
A.Accession: S76464
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-465 <KAN>
A.Cross-references: EMBL:D90915; GB:AE001339; NID:g1653604; PIDN:BA18593.1; PID:d101932
A.Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 465;
Matches 9; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 8 FVGOLFQGHDPVFLM 23
DB 25 WVGQIFSLQDLKFLYV 40

RESULT 38
S51613
C.Species: *Candida albicans*
C.Date: 15-Jul-1995 #sequence_revision 08-Sep-1995 #text_change 08-Sep-1995
C.Accession: S51613; S49207
R.Sherlock, G.; Bahman, A.M.; Mahal, A.; Shieh, J.C.; Ferreira, M.; Rosamond, J.
Mol. Gen. Genet. 245, 716-723, 1994
A.Title: Molecular cloning and analysis of CDC28 and cyclin homologues from the human fu
A.Reference number: S51611; MUID:95131949; PMID:7830719
A.Accession: S51613
A.Molecule type: DNA
A.Residues: 1-465 <SHE>
A.Cross-references: EMBL:X80033

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 465;
Matches 6; Conservative 8; Mismatches 2; Indels 0; Gaps 0;

QY 6 SSEVGLFGQHPDVFY 21
DB 211 SWIGELFOFIPNITY 226

RESULT 39
C83052
C.Species: *Pseudomonas aeruginosa*
C.Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C.Accession: C83052
R.Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,
.; Lorry, S.; Olson, M.V.
Nature 406, 959-964, 2000
A.Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pat
A.Reference number: A82950; MUID:20437337; PMID:10984043
A.Accession: C83052
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-701 <STO>
A.Cross-references: GB:AE004886; GB:AE004091; NID:g9951001; PIDN:AG08126.1; GSPDB:GN001
A.Experimental source: strain PA01
C.Genetics:
A.Gene: PNP; PA4740
C.Superfamily: polyribonucleotide nucleotidyltransferase alpha chain

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 701;
Matches 32.1%; Pred. No. 1.2e+02;

Matches 9; Conservative 6; Mismatches 9; Indels 4; Gaps 1;
QY 4 SGSSFVGQL----TGQHPDVFYLMPEPM 27
DB 140 SGIFPAGPIGAARVGFHPEIGYILNPTY 167

RESULT 40
D96521
C.Species: *Arabidopsis thaliana* (mouse-ear cress)
C.Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C.Accession: D96521
R.Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Comp, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A.Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lures, J.S.; Malt, R.; Marzia
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A.Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A.Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.
A.Reference number: A86141; MUID:21016719; PMID:11130712
A.Accession: D96521
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-1340 <STO>
A.Cross-references: GB:AE005173; NID:g8778513; PIDN:AAF9521.1; GSPDB:GN00141
C.Genetics:
A.Gene: F21D18.16
A.Map position: 1

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 1340;
Matches 10; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 9 VGLFGQHPDVFYLMPE 26
DB 658 VGDLHGQLHDLXYLMODA 675

RESULT 41
AB0480
C.Species: *Yersinia pestis*
C.Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 02-Nov-2001
C.Accession: AB0480
R.Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barre
Nature 413, 523-527, 2001
A.Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
A.Reference number: AB0001; MUID:21470413; PMID:11586360
A.Accession: AB0480
A.Status: preliminary
A.Molecule type: DNA
A.Residues: 1-3013 <KUR>
A.Cross-references: GB:AL590842; PIDN:CAC93406.1; PID:g15981852; GSPDB:GN00175
C.Genetics:
A.Gene: YPO3944

Query Match
Best Local Similarity 28.3%; Score 47; DB 2; Length 3013;
Matches 9; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

QY 1 SMRSSEVGLFGQHPDVFYLMPEPM 27
DB 2942 TWMGSKRAVGNLWSEWGDPSAYTAPGM 2968

RESULT 42
ESDKM

oleoyl-[acyl-carrier-protein] hydrolase (EC 3.1.2.14) - mallard
N:Alternate names: S-acyl fatty acid synthase, thioesterase
C:Species: Anas platyrhynchos (mallard)
C:Date: 28-Feb-1986 #sequence_revision 28-Feb-1986 #text_change 26-May-2000
C:Accession: A00775
R:Poulos, A.J.; Rogers, L.; Cheesbrough, T.M.; Kolattukudy, P.E.
J. Biol. Chem. 260, 15953-15958, 1985
A:Title: Cloning and sequencing of the cDNA for S-acyl fatty acid synthase thioesterase
A:Reference number: A00775; MUID:86059488; PMID:2415525
A:Accession: A00775
A:Molecule type: mRNA
A:Residues: 1-251 <PDB>
A:Cross-references: GB:M12101; NID:9213098; PIDN:AAA9222.1; PID:9213099
C:Comment: This protein, isolated from the uropygial gland, catalyzes the hydrolytic re-
fatty acid synthase.
C:Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-[acyl-carrier-protein] hy-
C:Keywords: fatty acid biosynthesis; thioester hydrolase
F:17-231/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 28.0%; Score 46.5; DB 1; Length 251;
Best Local Similarity 47.8%; Pred. No. 46;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 4 SSSSFVQLEFGHPDYFYLMEPA 26
DB 214 SGGTSTSLPGNH---FYLMEPS 233

RESULT 43
I50520
S-acyl fatty acid synthase thioesterase - duck
C:Species: Anas platyrhynchos (domestic duck)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 18-Jun-1999
C:Accession: I50520
R:Sasaki, G.C.; Cheesbrough, V.; Kolattukudy, P.E.
DNA 7, 449-457, 1988
A:Title: Nucleotide sequence of the S-acyl fatty acid synthase thioesterase gene and its
A:Reference number: I50520; MUID:89090803; PMID:2850144
A:Accession: I50520
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-251 <SAS>
A:Cross-references: GB:M21635; NID:9213092; PIDN:AAA9219.1; PID:9213093
C:Genetics:
A:Introns: 44/1; 90/2; 123/3; 180/2; 208/1
C:Superfamily: oleoyl-[acyl-carrier-protein] hydrolase; oleoyl-[acyl-carrier-protein] hy-
F:17-231/Domain: oleoyl-[acyl-carrier-protein] hydrolase homology <ACPH>

Query Match 28.0%; Score 46.5; DB 2; Length 251;
Best Local Similarity 47.8%; Pred. No. 46;
Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;

OY 4 SSSSFVQLEFGHPDYFYLMEPA 26
DB 214 SGGTSTSLPGNH---FYLMEPS 233

RESULT 44
A28550
cyclohexanone monooxygenase (EC 1.14.13.22) - Acetobacter sp.
C:Species: Acetobacter sp.
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 08-Oct-1999
C:Accession: A28550; B28550
R:Chen, Y.C.J.; Peoples, O.P.; Walsh, C.T.
J. Bacteriol. 170, 781-789, 1988
A:Title: Acetobacter cyclohexanone monooxygenase: gene cloning and sequence determinat
A:Reference number: A28550; MUID:88115180; PMID:3338974
A:Contents: NCBI 9871
A:Accession: A28550
A:Molecule type: DNA
A:Residues: 1-542 <CH1>
A:Cross-references: GB:M19029; NID:9141767; PIDN:AAA21892.1; PID:9141768
A:Accession: B28550

A:Molecule type: protein
A:Residues: 2-11 <CH2>
C:Comment: The initiator Met is not shown.
C:Keywords: oxidoreductase

Query Match 28.0%; Score 46.5; DB 2; Length 542;
Best Local Similarity 32.0%; Pred. No. 11e+02;
Matches 8; Conservative 8; Mismatches 8; Indels 1; Gaps 1;

OY 2 WRSG--SSFFVQLEFGHPDYFYLMEP 25
DB 403 WKFGPSYGVTVVNNPNNFWLGP 427

RESULT 45
F82623
potassium uptake protein XF1903 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 20-Aug-2000
C:Accession: F82623
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A>Note: for a complete list of authors see reference number A59328 below
A:Accession: F82623
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-634 <SIM>
A:Cross-references: GB:AE004010; GB:AE003849; NID:99106992; PIDN:AF84709.1; GSPDB:GN
R:Experimental source: strain 9a5c
A:Authors: J.G. Reinach, F.C. Arruda, P. Abreu, F.A. Acencio, M. Alvarenga, R.
Brites, M.R.S. Bueno, M.R.P. Camargo, A.A. Camargo, L.E.A. Carraro, D.M. Carer
as-Neco, E. Docena, C. El-Dorri, H. Facinanci, A.P. Ferreira, A.J.S.
submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr
J.D.; Junqueira, M.L.; Kemper, E.L.; Klatjima, J.P.; Krieger, J.E.; Kuramae, E.E.; La
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins
A:Authors: Martins, E.M.F.; Matsumura, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri,
Rodrigues, V.; Rosa, A.J.; de M. de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sava
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1903

Query Match 28.0%; Score 46.5; DB 2; Length 634;
Best Local Similarity 25.0%; Pred. No. 1.3e+02;
Matches 11; Conservative 6; Mismatches 8; Indels 19; Gaps 1;

OY 3 RSGSSFVQLEFGHPDYFYLMEPAW 27
DB 170 RGTFRVKGTFPGFTLLMFIAIGVGVYINIAQAPVLAIPNSW 213

RESULT 46
H86246
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: H86246
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dedar
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Matzla
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallo
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: H86246
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1820 <SNO>
A:Cross-references: GB:AE005172; NID:q5734728; PIDN:AD49993.1; GSPDB:GN00141
C:Genetics:
A:Map position: 1
C:Superfamily: S-receptor kinase; protein kinase homology; S-locus-specific glycoprotein

Query Match 28.0%; Score 46.5; DB 2; Length 820;
Best Local Similarity 52.6%; Pred. No. 1.7e+02;
Matches 10; Conservative 3; Mismatches 5; Indels 1; Gaps 1;

QY 2 WRSQSFVQGLFGHPDYF 20
DB 215 WRSQSFVQGLFGHPDYF 232

RESULT 47
T08948
Hypothetical protein F25024.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jan-2000
C:Accession: T08948
R:Bayan, M.; Hilbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Bancroft, submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16519
A:Accession: T08948
A:Molecule type: DNA
A:Residues: 1-175 <BEV>
A:Cross-references: EMBL:AL078469; GSPDB:GN00062; ATSP:F25024.50
A:Experimental source: cultivar Columbia; BAC clone F25024
C:Genetics:
A:Gene: ATSP:F25024.50
A:Map position: 4
A:Introns: 167/3
C:Superfamily: Arabidopsis hypothetical protein F25G13.20

Query Match 27.7%; Score 46; DB 2; Length 175;
Best Local Similarity 50.0%; Pred. No. 36;
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFG 14
DB 97 TWRKGFSPAGNPYG 110

RESULT 48
T46754
AdcC protein [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 19-May-2000
C:Accession: T46754
R:Claverys, J.P.Y.
Submitted to the EMBL Data Library, October 1998
A:Reference number: Z24090
A:Accession: T46754
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-234 <CLAA>
A:Cross-references: EMBL:Z71552; PIDN:CA96186.1
A:Experimental source: strain R6
C:Genetics:
A:Gene: adcC
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 27.7%; Score 46; DB 2; Length 234;
Best Local Similarity 47.1%; Pred. No. 50;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 VGOIFGHPDYFLMEP 25
DB 148 IARMFASDPDYFLMEP 164

Db 148 IARMFASDPDYFLMEP 164

RESULT 49
F98118
Hypothetical protein adcC [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 17-May-2002
C:Accession: F98118
R:Hoskins, J.A.; Albom, J.R.; Arnold, J.; Blaszczyk, L.; Bueget, S.; Dehoff, B.S.; e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McHenry, S.; y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaekunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: F98118
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:AE007317; PIDN:AAL00779.1; PID:q15459675; GSPDB:GN00174
C:Genetics:
A:Gene: adcC
C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

Query Match 27.7%; Score 46; DB 2; Length 234;
Best Local Similarity 47.1%; Pred. No. 50;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 VGOIFGHPDYFLMEP 25
DB 148 IARMFASDPDYFLMEP 164

RESULT 50
H95253
Zinc ABC transporter, ATP-binding protein SP2171 [imported] - Streptococcus pneumoniae
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: H95253
R:Letellin, H.; Nelson, K.E.; Paulsen, I.T.; Bisen, J.A.; Read, T.D.; Peterson, S.; H on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf l nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: H95253
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-234 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK76225.1; PID:q14973683; GSPDB:GN00164; TIGR: A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP2171

Query Match 27.7%; Score 46; DB 2; Length 234;
Best Local Similarity 47.1%; Pred. No. 50;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 9 VGOIFGHPDYFLMEP 25
DB 148 IARMFASDPDYFLMEP 164

Search completed: February 20, 2003, 13:34:22
Job time : 72.4 secs

983 35.5 21.4 1177 1 TAK DROME 024592 drosophila
 984 35.5 21.4 1197 1 EVGS_ECO57 P8402 escherichia
 985 35.5 21.4 1197 1 EVGS_ECOLI P0985 escherichia
 986 35.5 21.4 1316 1 RPOC_MCTV P47769 mycobacteri
 987 35.5 21.4 1612 1 RPOC_PMVU P99657 pepper mild
 988 35.5 21.4 1612 1 RPOC_PAMV P99098 pepper mild
 989 35.5 21.4 1616 1 RPOC_TOMK2 P96676 tomato mosa
 990 35.5 21.4 1616 1 RPOC_TOML P03587 tomato mosa
 991 35.5 21.4 1616 1 RPOC_TOMK1 P5870 tomato mosa
 992 35.5 21.4 1694 1 CLH_DICD1 Q12397 dicystostell
 993 35.5 21.4 2181 1 STCA_EMBE1 P26661 h genome po
 994 35.5 21.4 3033 1 POLG_HCVJ8 P27742 emericella
 995 35.5 21.4 3770 1 ACVS_EMBE1 P27742 emericella
 996 35.5 21.4 4377 1 ANK3_HUMAN Q12955 homo sapien
 997 35.5 21.4 4590 1 PATH_HUMAN Q14517 homo sapien
 998 35 21.1 104 1 LAC2_RAT P20767 rattus norv
 999 35 21.1 114 1 YG13_BACTU P10024 bacillus th
 1000 35 21.1 128 1 OTOR_MOUSE Q9163 mus musculu

ALIGNMENTS

RESULT 1
 C6ST_CHICK STANDARD: PRT: 458 AA.
 ID C6ST_CHICK
 AC Q92179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Chondroitin 6-sulfotransferase (EC 2.8.2.17) (C6ST).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-44; 46-52 AND 186-192.
 RC STRAIN=White Leghorn; TISSUE=Embryonic chondrocytes;
 RX MEDLINE=95355490; PubMed=7629189;
 RA Fukuta M., Uchimura K., Nakashima K., Kato M., Kimata K.,
 RA Shiomura T., Habuchi O.;
 RT "Molecular cloning and expression of chick chondrocyte chondroitin 6-
 sulfotransferase";
 RL J. Biol. Chem. 270:18575-18580(1995).
 CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN
 SULFATE.
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + chondroitin =
 CC adenosine 3',5'-bisphosphate + chondroitin 6'-sulfate.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE
 CC (BY SIMILARITY).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: D49915; BAA08655.1; -
 DR Transferrase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
 FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 24 37 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DOMAIN 38 458 LUMENAL, CATALYTIC (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 226 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 339 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 458 AA: 52253 MW: 69A3B7D0A5086F0C CRC64;

Query Match 60.8%; Score 101; DB 1; Length 458;
 Best Local Similarity 59.3%; Pred. No. 3.5e-07;
 Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFTVQGLFGHPDYFIMEPAMHY 29
 Db 122 RTGSFVEGFEPFNOGNIFLFEPLMHI 148

RESULT 2

ARGD_METHH STANDARD: PRT: 390 AA.
 ID ARGD_METHH
 AC Q27392;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Acetylornithine aminotransferase (EC 2.6.1.11) (ACOAT).
 OS Methanobacterium thermoautotrophicum.
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
 OC Methanobacteriaceae; Methanothermobacter.
 OX NCBI_TaxID=187420;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Delta H;
 RX MEDLINE=98037514; PubMed=9371463;
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
 RA Aldridge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
 RA Spadafora R., Viacre R., Wang Y., Wierzbowski J., Gibson R.,
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,
 RA McDougall S., Shiner G., Goyal A., Pietrowski S., Church G.M.,
 RA Daniels C.J., Mao J.-F., Rice P., Noelling J., Reeve J.N.;
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum
 deltaH: functional analysis and comparative genomics";
 RL J. Bacteriol. 179:7135-7155(1997).
 CC -1- CATALYTIC ACTIVITY: N(2)-acetyl-L-ornithine + 2-oxoglutarate = N-
 CC acetyl-L-glutamate 5-semialdehyde + L-glutamate.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: Arginine biosynthesis; fourth step.
 CC -1- SIMILARITY: BELONGS TO CLASS-III OF PYRIDOXAL-PHOSPHATE-DEPENDENT
 CC AMINOTRANSFERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE00897; AAB85815.1; -
 DR HSSP: P12985; 10J3
 DR InterPro: IPR000954; Aminotran_3.
 DR InterPro: IPR004636; ArgD.
 DR Pfam: PF00202; aminotran_3; 1.
 DR TIGRfam: TIGR00707; argD; 1.
 DR PROSITE: PS00600; AA_TRANSFER_CLASS_3; 1.
 DR TRANSFERRASE; Aminotransferase; Pyridoxal phosphate;
 KW Arginine biosynthesis; Complete proteome
 FT BINDING 246 246 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT SEQUENCE 390 AA: 41763 MW: 5F309E073D7DA8FC CRC64;
 SQ

RESULT 3

Query Match 31.9%; Score 53; DB 1; Length 390;
 Best Local Similarity 52.9%; Pred. No. 2.8;
 Matches 9; Conservative 3; Mismatches 5; Indels 0; Gaps 0;
 QY 3 RSGSFTVQGLFGHPDY 19
 Db 225 RTGAFASOLFVGPDI 241

```

NOEF_RHISN
ID NOEF_RHISN STANDARD: PRT: 419 AA.
AC P55472: P72326:
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-DEC-1998 (Rel. 37, Last annotation update)
DE Modulation protein noe (EC 2.8.2.-).
GN NOEF OR Y4HB.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234.
OC Bacteria: Proteobacteria: alpha subdivision: Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-97361801: PubMed-9218762;
RA Hann M., Jabouri S., Quesada-Vincens S., Freilberg C., Perret X.,
RA Broughton W.J., Fellay R.,
RT Sulfation of Rhizobium sp. NGR234 nod factors is dependent on noe,
a new host-specificity gene.
RL Mol. Microbiol. 24:1119-1129(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-97305956: PubMed-9163424;
RA Freilberg C.A., Fellay R., Baitoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- FUNCTION: REQUIRED FOR THE FORMATION OF SULFATED NOD FACTOR.
CC PROPOSED TO TRANSFER ACTIVATED SULFATE (PAPS) TO THE FUCOSE
CC OF THE NOD FACTOR.
CC -1- SIMILARITY: LIMITED TO NODH AND TO C. ELEGANS F42G9.8.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Y09415; CAA70569.1; -
CC DR EMBL: AE000076; AAB91690.1; -
CC KM Modulation; Transferrase; Plasmid.
SQ SEQUENCE 419 AA; 46569 MW; 84BC48B0416AAAF CRC64;

Query Match 31.6%; Score 52.5; DB 1; Length 419;
Best Local Similarity 44.0%; Pred. No. 3.5;
Matches 11; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 3 RSGSFGVGFQGHDPVFLMEPAW 27
DB 19 RSGTTLAHLQGHDPITAPPEP-W 42

```

```

RESULT 4
MNTB_LISIN
ID MNTB_LISIN STANDARD: PRT: 240 AA.
AC C92AF9;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Manganese transport system ATP-binding protein mntB.
DE Manganese transport system ATP-binding protein mntB.
GN MNTB OR LIN1963.
OS Listeria innocua.
OC Bacteria: Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-CLIP 11262 / Serovar 6a;
RX MEDLINE-21537279: PubMed-11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Vazquez-Boland J.-A., Boucher P., Bloeker H., Brandt P., Chakraborty T.,
Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,

Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hant J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.,
RT "Comparative genomics of Listeria species.";
RL Science 294:849-852(2001).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system. Probably responsible for energy coupling to the transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL596170; CAC97193.1; -
CC DR Listlist; LIN01963; -
CC DR InterPro: IPR003439; ABC_transporter.
CC Pfam: PF00005; ABC_tran; 1
CC DR Prodom: PD000006; ABC_transporter; 1.
CC DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC KW Transport; ATP-binding; Membrane; Complete proteome.
FT NP_BIND 33 40 ATP (POTENTIAL).
SQ SEQUENCE 240 AA; 26470 MW; ADD256CE8AC20B94 CRC64;

Query Match 28.9%; Score 48; DB 1; Length 240;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 8 FVGQIFGQHPVFLMEP 25
DB 143 FVARLAQHARFLDEP 160

```

```

RESULT 5
MNTB_LISMO
ID MNTB_LISMO STANDARD: PRT: 240 AA.
AC Q8Y651;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Manganese transport system ATP-binding protein mntB.
DE Manganese transport system ATP-binding protein mntB.
GN MNTB OR LM01849.
OS Listeria monocytogenes.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1639;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-EGD-e / Serovar 1/2a;
RX MEDLINE-21537279: PubMed-11679669;
RA Glaser P., Frangoul L., Buchrieser C., Rusniok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Charbit A., Chetouani F., Couve E., de Daruvar A., Deloux P.,
RA Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Entian K.-D., Fsihi H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hant J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkak G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordstiek G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tierrez A.,
RA Vazquez-Boland J.-A., Voss H., Wehlend J., Cossart P.,
RT "Comparative genomics of Listeria species.";

```



```

RL Science 294:849-852(2001).
CC -1- FUNCTION: This protein is probably a component of a manganese
CC permease, a binding protein-dependent, ATP-driven transport
CC system. Probably responsible for energy coupling to the transport
CC system (By similarity).
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591981; CAC99927.1; -.
CC Listlist: LMO01849; -.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR003439; ABC_transportr.
CC Pfam: PF00005; ABC_tran.1.
CC ProDom: PD000006; ABC_transportr; 1.
CC SMART: SM00382; AAA; 1.
CC PROSITE: PS00211; ABC_TRANSPORTER; 1.
CC Transport: ATP-binding; Membrane; Complete proteome.
CC NP_BIND: 33 40
CC FT SEQUENCE 240 AA; 26495 MW; C3667E78F4D853E4 CRC64;
SQ
Query Match 28.9%; Score 48; DB 1; Length 240;
Best Local Similarity 38.9%; Pred. No. 8.8;
Matches 7; Conservative 5; Mismatches 6; Indels 0; Gaps 0;
OY 8 FVGOLFQHPDYFYMPEP 25
Db 143 FTRALAHAEITFIDEP 160
-----
RESULT 6
GATE_MYCPN STANDARD; PRT: 338 AA.
ID GATE_MYCPN STANDARD; PRT: 338 AA.
AC P75517;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE UDP-glucose 4-epimerase (EC 5.1.3.2) (Galactowaldenase) (UDP-
DE galactose 4-epimerase).
GN GALE OR MPN257 OR MP576.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelfeich R., Hilbert H., Plagens H., Pirk E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae."
RT Nucleic Acids Res. 24:4420-4449(1996).
CC -1- CATALYTIC ACTIVITY: UDP-glucose = UDP-galactose.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: Galactose metabolism; third step.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: AE000056; AAB96224.1; -.
DR HSSP: P09147; 1KVS.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
DR TIGRFAMs: TIGR01179; gale; 1.
KW Isomerase; NAD; Galactose metabolism; Complete proteome.
FT NP_BIND: 38
FT SEQUENCE 338 AA; 38132 MW; 9C50FF385B568C03 CRC64;
SQ
Query Match 28.9%; Score 48; DB 1; Length 338;
Best Local Similarity 60.0%; Pred. No. 13;
Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;
OY 5 GGSFVGOLFQHPDY 19
Db 18 GSCFIDQLKQHPDY 32
-----
RESULT 7
GATA_AOUAE STANDARD; PRT: 478 AA.
ID GATA_AOUAE STANDARD; PRT: 478 AA.
AC O66610;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE GlutamyI-tRNA(Gln) amidotransferase subunit A (EC 6.3.5.-) (Glu-ADT
DE subunit A).
GN GATA OR AQ_247.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
OC Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VFS;
RX MEDLINE=98196666; PubMed=9537320;
RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Aujay M., Huber R.,
RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
RT "The complete genome of the hyperthermophilic bacterium Aquifex
RT aeolicus."
RL Nature 392:353-358(1998).
CC -1- FUNCTION: FURNISHES A MEANS FOR FORMATION OF CORRECTLY CHARGED
CC GIN-tRNA(GIN) THROUGH THE TRANSAMIDATION OF MISACTYLATED GLU-
CC TENA(GIN) IN ORGANISMS WHICH LACK GLUTAMINYI-TRNA SYNTHETASE. THE
CC REACTION TAKES PLACE IN THE PRESENCE OF GLUTAMINE AND ATP THROUGH
CC AN ACTIVATED GAMMA-PHOSPHO-GLU-tRNA(GIN) (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamyI-tRNA(Gin) + L-glutamine = ADP
CC + phosphate + L-glutaminyI-tRNA(Gin) + L-glutamate.
CC -1- SUBUNIT: HETERODIMER OF A, B AND C SUBUNITS (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE AMIDASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AE000680; AAC06569.1; -.
CC InterPro: IPR000120; Amidase.
CC InterPro: IPR004412; Gata.
CC Pfam: PF01425; Amidase; 1.
CC TIGRFAMs: TIGR00132; gata; 1.
CC PROSITE: PS00571; AMIDASES; 1.
KW Protein biosynthesis; Ligase; Complete proteome.
KW SEQUENCE 478 AA; 53524 MW; 72EB38259D60B6 CRC64;
SQ
Query Match 28.9%; Score 48; DB 1; Length 478;
Best Local Similarity 36.7%; Pred. No. 18;
Matches 11; Conservative 4; Mismatches 9; Indels 6; Gaps 1;

```



```

FT DOMAIN 1556 1753 LAMININ G-LIKE 1.
FT DOMAIN 1756 1792 EGF-LIKE 2 CALCIUM-BINDING.
FT DOMAIN 1796 1963 LAMININ G-LIKE 2.
FT DOMAIN 1965 2000 EGF-LIKE 3 CALCIUM-BINDING.
FT DOMAIN 2091 2126 EGF-LIKE 4 CALCIUM-BINDING.
FT DOMAIN 2744 2802 GFS.
FT DOMAIN 140 143 POLY-HIS.
FT DOMAIN 155 159 POLY-ARG.
FT DOMAIN 2567 2579 POLY-SER.
FT DOMAIN 3460 3467 POLY-GLN.
FT DISULFID 1486 1497 POTENTIAL.
FT DISULFID 1491 1506 POTENTIAL.
FT DISULFID 1508 1517 POTENTIAL.
FT DISULFID 1760 1771 POTENTIAL.
FT DISULFID 1765 1780 POTENTIAL.
FT DISULFID 1782 1791 POTENTIAL.
FT DISULFID 1969 1979 POTENTIAL.
FT DISULFID 1973 1988 POTENTIAL.
FT DISULFID 1990 1999 POTENTIAL.
FT DISULFID 2092 2095 POTENTIAL.
FT DISULFID 2097 2114 POTENTIAL.
FT DISULFID 2116 2125 POTENTIAL.
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 179 179 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 886 886 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

Query Match 28.9%; Score 48; DB 1; Length 3579;
Best Local Similarity 42.9%; Pred. No. 1.5e+02;
Matches 9; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```

```

OY 4 GSSSFVQLFGQHPDPVYIME 24
Db 3374 GSGSOTIGHMSEFHPDAAYLSD 3394

```

```

RESULT 11
Y4IE_RHISN STANDARD; PRT; 135 AA.
AC P55488;

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE Hypothetical 15.4 kDa protein Y4IE.
GN Y4IE.

```

```

OS Rhizobium sp. (strain NGR234).
OG Plasmid sym pNGR234.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=394;

```

```

RN RP
RP MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Balroch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- SIMILARITY: NONE OBVIOUS OUTSIDE OF SIMILARITY WITH Y4RL AND Y4ZA.
CC -1- SIMILARITY: NONE OBVIOUS OUTSIDE OF SIMILARITY WITH Y4RL AND Y4ZA.
CC POTENTIAL FRAGMENT.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

```

DR EMBL; AE000078; AAB91700.1; -
KM Hypothetical protein; Plasmid;
SQ SEQUENCE 135 AA; 15402 MW; 8166C909058F14C3 CRC64;

```

```

Query Match 28.3%; Score 47; DB 1; Length 135;
Best Local Similarity 43.8%; Pred. No. 6.7;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 2 WRSGSFVQLFGQHP 17
Db 76 WRTGSEFLSLQAEYP 91

```

```

RESULT 12
ID RMLB_STRMU STANDARD; PRT; 348 AA.
AC P95780;

```

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE dTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
GN RMLB.

```

```

OS Streptococcus mutans.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1309;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-XC;
RA Tsukiyoka Y., Yamashita Y., Nakano Y., Oho T., Koga T.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
CC glucose + H(2)O.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: dTDP-L-thiamose biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTPD-GLUCOSE
CC DEHYDRATASE SUBFAMILY.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).

```

```

CC EMBL: D78182; BAA11249.1; -
DR HSSP: P27830; 1BXK.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase_1.
DR TIGRfams: TIGR01181; dTDP-gluc_dehyd; 1.
KM Lipopolysaccharide biosynthesis; Lyase; NAD.
FT NP_BIND 11 17 NAD (POTENTIAL).
FT SEQUENCE 348 AA; 39251 MW; 034E68BEB15079DD CRC64;

```

```

Query Match 28.3%; Score 47; DB 1; Length 348;
Best Local Similarity 53.3%; Pred. No. 18;
Matches 8; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

```

```

OY 5 GSSSFVQLFGQHPDV 19
Db 17 GSNFVHYVNNHPDV 31

```

```

RESULT 13
ID CG12_CANAL STANDARD; PRT; 465 AA.
AC P43062;

```

```

DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DE G1/S-specific cyclin CLN2.
GN CLN2.

```

```

OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;

```


CC USUALLY C16, HOWEVER IN THE MAMMARY GLANDS OF NONRUMINANT MAMMALS,
 CC AND IN THE UROGYAL GLAND OF CERTAIN WATERFOWL EXISTS A SECOND
 CC THIOESTERASE WHICH RELEASE MEDIUM-CHAIN LENGTH FATTY ACIDS (C8 TO
 CC C22).
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl]-
 CC carrier protein] + oleate.
 CC -1- SIMILARITY: TO OTHER THIOESTERASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M12101; AAA49222.1; -;
 CC DR EMBL; M21635; AAA49219.1; -;
 CC DR PIR; A00775; ESDKTM.
 CC DR InterPro: IPR000379; Ser. esters. site.
 CC DR InterPro: IPR001031; Thioesterase.
 CC DR Pfam: PF00975; Thioesterase; 1.
 CC KW Fatty acid biosynthesis; Hydrolyase.
 CC FT ACT_SITE 90 90 PROBABLE.
 CC FT ACT_SITE 226 226 BY SIMILARITY.
 CC FT CONFLICT 43 43 V -> E (IN REF. 2).
 CC SQ SEQUENCE 251 AA; 28807 MW; 3DFA25170D9EDC50 CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 251;
 CC Best Local Similarity 47.8%; Pred. No. 15;
 CC Matches 11; Conservative 3; Mismatches 6; Indels 3; Gaps 1;
 CC
 CC QY 4 SSGSSEVGLFGQHPDVFYIMPPA 26
 CC DB 214 SGTSTSTSLPQNH--FYLMPPS 233
 CC
 CC RESULT 16
 CC CARD_ERWCA STANDARD; PRT; 376 AA.
 CC ID CARD_ERWCA
 CC AC 09XB58;
 CC DT 30-MAY-2000 (Rel. 39, Created)
 CC DT 30-MAY-2000 (Rel. 39, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Carabapenam antibiotics biosynthesis protein CARD.
 CC GN CARD.
 CC OS Erwina carotovora.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC OC Pectobacterium.
 CC NCBI_TaxID=354;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=ATCC 39048 / GS101;
 CC RX MEDLINE=9806591; PubMed=9402024;
 CC RA McGowan S.J., Sebahia M., O'Leary S., Hardie K.R., Williams P.,
 CC Stewart G.S., Bycroft B.W., Salmon G.P.,
 CC "Analysis of the carabapenam gene cluster of Erwina carotovora:
 CC RT definition of the antibiotic biosynthetic genes and evidence for a
 CC novel beta-lactam resistance mechanism.";
 CC RL Mol. Microbiol. 26:545-556(1997).
 CC -1- PATHWAY: Carabapenam antibiotics biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE PROLINE OXIDASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; U17224; AAD38232.1; -;
 CC DR InterPro: IPR002872; Pro.dh.

DR Pfam: PF01619; Pro.dh: 1.
 KW Antibiotic biosynthesis; Oxidoreductase.
 SQ SEQUENCE 376 AA; 42885 MW; 1F8AF5AET7330E7C CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 376;
 CC Best Local Similarity 36.7%; Pred. No. 23;
 CC Matches 11; Conservative 8; Mismatches 10; Indels 1; Gaps 1;
 CC
 CC QY 1 SRRSGSFFVGLFGQHPDVFY-LMEPAMVY 29
 CC DB 298 SLRAGRTWGLYGLDHTTSLQGFVY 327
 CC
 CC RESULT 17
 CC CRYO_ACISP STANDARD; PRT; 542 AA.
 CC ID CRYO_ACISP
 CC AC P12015;
 CC DT 01-OCT-1989 (Rel. 12, Created)
 CC DT 01-OCT-1989 (Rel. 12, Last sequence update)
 CC DT 01-AUG-1991 (Rel. 19, Last annotation update)
 CC DE Cyclohexanone monooxygenase (EC 1.14.13.22).
 CC OS Acinetobacter sp.
 CC OC Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
 CC OC Acinetobacter.
 CC NCBI_TaxID=472;
 CC RN [1]
 CC RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-11.
 CC RC STRAIN=NCIB 9871;
 CC RX MEDLINE=88115180; PubMed=3338974;
 CC RA Chen Y.-C.C., Peoples O.P., Walsh C.T.;
 CC "Acinetobacter cyclohexanone monooxygenase: gene cloning and sequence
 CC RT determination.";
 CC RL J. Bacteriol. 170:781-789(1988).
 CC -1- CATALYTIC ACTIVITY: Cyclohexanone + NADPH + O(2) = 6-hexanolide +
 CC NADP(+) + H(2)O.
 CC -1- COFACTOR: FAD.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC DR EMBL; M19029; AAA21892.1; -;
 CC DR PIR; A28550; A28550.
 CC DR InterPro: IPR000960; Flav.conc._mnoxgn.
 CC DR Pfam: PF00743; FMO-Like; 1.
 CC KW Aromatic hydrocarbons catabolism; Oxidoreductase; Monooxygenase;
 CC NADP; Flavoprotein; FAD.
 CC FT INIT_MET 0 0
 CC FT NP_BIND 6 18 FAD (BY SIMILARITY).
 CC FT NP_BIND 176 208 NADP (BY SIMILARITY).
 CC SQ SEQUENCE 542 AA; 60760 MW; 7ACDD25C7BD388F CRC64;
 CC
 CC Query Match 28.0%; Score 46.5; DB 1; Length 542;
 CC Best Local Similarity 32.0%; Pred. No. 34;
 CC Matches 8; Conservative 8; Mismatches 8; Indels 1; Gaps 1;
 CC
 CC QY 2 WRSGLSFFVGLFGQHPDVFYIMPP 25
 CC DB 403 WKBPSSYMGVYVNNYPMFMVILGP 427
 CC
 CC RESULT 18
 CC ADCC_STRPN STANDARD; PRT; 234 AA.
 CC ID ADCC_STRPN
 CC AC 087862;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 41, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Zinc transport system ATP-binding protein adcc.
 CC

GN ADCC OR SP2171.
 OS Streptococcus pneumoniae.
 OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 OC Streptococcus
 NCBI_TaxID=1313;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-R6;
 RX MEDLINE=98438930; PubMed=9765793;
 RA Dintilhac A., Claviers J.-P.;
 RT "The *adc* locus, which affects competence for genetic transformation in
 RT Streptococcus pneumoniae, encodes an ABC transporter with a putative
 RT lipoprotein homologous to a family of streptococcal adhesins.";
 RL Res. Microbiol. 148:119-131(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-TIGR4;
 RX MEDLINE=21357209; PubMed=11463916;
 RA Tettelin H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,
 RA Peterson S., Heidelberg J., Deboy R.T., Hatt D.H., Dodson R.J.,
 RA Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,
 RA Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,
 RA Holtapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,
 RA McDonald L.A., Feldblyum T.V., Angluoni S., Dickson T., Hickey E.K.,
 RA Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C., Hickey E.K.,
 RA Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;
 RT "Complete genome sequence of a virulent isolate of Streptococcus
 RT pneumoniae.";
 RL Science 293:498-506(2001).
 RN [3]
 RP FUNCTION.
 RC MEDLINE=98025470; PubMed=9379902;
 RA Dintilhac A., Allouin G., Granadel C., Claviers J.-P.;
 RT "Competence and virulence of Streptococcus pneumoniae: *Adc* and *PSa*
 RT mutants exhibit a requirement for Zn and Mn resulting from
 RT inactivation of putative ABC metal permeases.";
 RL Mol. Microbiol. 25:727-739(1997).
 CC -1- FUNCTION: PART OF THE ATP-DRIVEN TRANSPORT SYSTEM ADABC FOR ZINC.
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: 271552; CAA96186.1;
 DR EMBL; AE007505; AAK76225.1;
 DR TIGR: SP2171;
 DR InterPro: IPR003593; AAA_ATPase.
 DR InterPro: IPR003439; ABC_transportr.
 DR Pfam: PF00005; ABC_tran.1.
 DR ProDom: PD000006; ABC_transportr.1.
 DR SMART: SM00382; AAA.1.
 DR PROSITE: PS00211; ABC_TRANSPORTER.1.
 DR ATP-binding; transport; zinc transport; Complete proteome.
 SQ SEQUENCE 234 AA; 26542 MW; 2AF782EF93D5C1A0 CRC64;

Query Match 27.7%; Score 46; DB 1; Length 234;
 Best Local Similarity 47.1%; Pred. No. 17;
 Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 9 VGOIFGQHPDYFYLMEP 25
 : : : | | | | |
 Db 148 IARFASDPDYFLDEP 164

RESULT 19
 CYAB_LEIDO
 ID CYAB_LEIDO STANDARD; PRT; 1331 AA.
 AC Q25263;

DT 15-JUL-1998 (Rel. 35, Created)
 DT 15-JUL-1998 (Rel. 35, Last sequence update)
 DT 15-DEC-1998 (Rel. 37, Last annotation update)
 DE Receptor-type adenylyl cyclase B (EC 4.6.1.1) (ATP pyrophosphate-
 DE lyase) (Adenylyl cyclase).
 GN RAC-B.
 OS Leishmania donovani.
 OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
 NCBI_TaxID=5661;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-IS Sudanese;
 RX MEDLINE=95340554; PubMed=7615561;
 RA Sanchez M.A., Zeoli D., Klamo E.M., Kavanaugh M.P., Landfear S.M.;
 RT "A family of putative receptor-adenylyl cyclases from Leishmania
 RT donovani.";
 RL J. Biol. Chem. 270:17551-17558(1995).
 CC -1- FUNCTION: COULD ACT AS A RECEPTOR FOR A UNKNOWN LIGAND.
 CC -1- CATALYTIC ACTIVITY: ATP = 3',5'-cyclic AMP + diphosphate.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN THE INSECT STAGE (PROMASTIGOTE)
 CC BUT NOT IN THE MAMMALIAN HOST STAGE OF THE PARASITE LIFE CYCLE.
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-3 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: U17043; AAT74999.1;
 DR InterPro: IPR001054; G_cyclase.
 DR Pfam: PF00211; guanylate_cyc.1.
 DR SMART: SM00044; CYCC.1.
 DR PROSITE: PS50125; GUANYLATE_CYCLASES_2.1.
 KW lase; cAMP synthesis; Transmembrane; Receptor; Glycoprotein.
 FT DOMAIN 1 33
 FT TRANSEM 34 898
 FT DOMAIN 35 919
 FT TRANSEM 899 919
 FT DOVAIN 920 1331
 FT CARBOHYD 255 255
 FT CARBOHYD 429 429
 FT CARBOHYD 558 558
 FT CARBOHYD 574 574
 FT CARBOHYD 657 657
 FT SEQUENCE 1331 AA; 144162 MW; 5FC3AA22DC1E5072 CRC64;

Query Match 27.7%; Score 46; DB 1; Length 1331;
 Best Local Similarity 36.4%; Pred. No. 1e+02;
 Matches 8; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 6 SSEVGOIFGQHPDYFYLMEP 27
 : : : | | | | |
 Db 390 AAYVNNFNEHPCVTQLMIGW 411

RESULT 20
 AATM_CHICK
 ID AATM_CHICK STANDARD; PRT; 423 AA.
 AC P00508;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
 DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
 OS Gallus gallus (Chicken).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 CC Gallus.
 NCBI_TaxID=9031;

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=66059504; PubMed=3840803;
RA Jausi R., Cotton B., Juretic N., Christen P., Schumperli D.;
RT "The primary structure of the precursor of chicken mitochondrial
RT aspartate aminotransferase. Cloning and sequence analysis of cDNA.";
RL J. Biol. Chem. 260:16060-16063(1985).
RN [2]
RP SEQUENCE OF 23-423.
RX MEDLINE=83238520; PubMed=6345546;
RA Graf-Hausner U., Wilson K.J., Christen P.;
RT "The covalent structure of mitochondrial aspartate aminotransferase
RT from chicken. Identification of segments of the polypeptide chain
RT invariant specifically in the mitochondrial isoenzyme.";
RL J. Biol. Chem. 258:8813-8826(1983).
RN [3]
RP GENE STRUCTURE.
RC STRAIN=White leghorn;
RX MEDLINE=90382432; PubMed=2401287;
RA Juretic N., Matthes U., Zlak M., Christen P., Jausi R.;
RT "Structure of the genes of two homologous intracellularly heterocytic
RT isoenzymes. Cytosolic and mitochondrial aspartate aminotransferase of
RT chicken.";
RL Eur. J. Biochem. 192:119-126(1990).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=92277655; PubMed=1593633;
RA MCPhealen C.A., Vincent M.G., Jansonsius J.N.;
RT "X-ray structure refinement and comparison of three forms of
RT mitochondrial aspartate aminotransferase.";
RL J. Mol. Biol. 225:495-517(1992).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL, M12105; AAA48603.1; -.
DR PIR, A00597; XNCHDM.
DR PIR, A24554; A24554.
DR PIR, S11346; S11346.
DR PDB, 7AAT; 31-OCT-93.
DR PDB, 8AAT; 31-OCT-93.
DR PDB, 9AAT; 31-OCT-93.
DR PDB, 1AKA; 31-JUL-94.
DR PDB, 1AKB; 31-JUL-94.
DR PDB, 1AKC; 31-JUL-94.
DR PDB, 1AKD; 31-JUL-94.
DR PDB, 1AMA; 31-OCT-93.
DR PDB, 1MAA; 31-OCT-93.
DR PDB, 1MAO; 31-JAN-94.
DR PDB, 1TAR; 31-JAN-94.
DR PDB, 1TAS; 31-JAN-94.
DR PDB, 1TAX; 31-JAN-94.
DR PDB, 1OXO; 10-JUN-96.
DR PDB, 1OXP; 10-JUN-96.
DR PDB, 1TVR; 23-JUL-97.
DR InterPro: IPR004839; AminoTransfl/2.
DR InterPro: IPR000796; AsptTransf_Sub.
DR InterPro: IPR004838; NHTransf_1.
DR Pfam: PF00155; aminotran_1_2; 1.
DR PRINTS; PR00799; TRANSAMINASE.

```

```

DR PROSITE; PS00105; AA_TRANSFER_CLASS_1; 1.
KW Transf. peptidase; AminoTransferase; Pyridoxal phosphate; Mitochondrion;
KM TRANSIT 1 22
FT CHAIN 23 423
FT BINDING 272 272
FT CONFLICT 67 67
FT CONFLICT 168 168
FT CONFLICT 216 216
FT TURN 27 28
FT TURN 36 37
FT HELIX 46 46
FT TURN 50 51
FT STRAND 53 54
FT STRAND 61 62
FT TURN 63 65
FT STRAND 66 67
FT HELIX 71 83
FT TURN 84 84
FT TURN 92 93
FT HELIX 96 107
FT HELIX 112 116
FT TURN 117 117
FT STRAND 119 125
FT HELIX 126 141
FT STRAND 147 152
FT TURN 156 157
FT HELIX 158 164
FT TURN 165 166
FT STRAND 168 173
FT STRAND 175 176
FT TURN 177 180
FT STRAND 181 182
FT HELIX 184 182
FT TURN 193 193
FT TURN 196 197
FT STRAND 199 203
FT TURN 208 210
FT HELIX 216 229
FT TURN 230 230
FT STRAND 232 237
FT TURN 240 241
FT HELIX 247 250
FT TURN 251 251
FT HELIX 252 259
FT TURN 260 261
FT STRAND 265 269
FT TURN 271 274
FT HELIX 277 279
FT STRAND 281 287
FT HELIX 291 309
FT HELIX 315 324
FT TURN 325 325
FT HELIX 327 358
FT TURN 359 359
FT HELIX 365 369
FT STRAND 374 376
FT HELIX 381 391
FT STRAND 393 394
FT STRAND 396 396
FT TURN 397 399
FT STRAND 400 402
FT HELIX 403 405
FT HELIX 408 410
FT TURN 411 422
SQ SEQUENCE 423 AA; 47241 MW; 59D65D4EDADD8B8A CRC64;
Query Match 27.4%; Score 45.5; DB 1; Length 423;
Best Local Similarity 40.7%; Pred. No. 37;
Matches 11; Conservative 6; Mismatches 9; Indels 1; Gaps 1;
QY 1 SWRGSFVQGLFGQHPDVFYLMPEAW 27
| | | | | : : : | | | | | : | : |

```

DB 130 SLRVGANEZQRFKFSRDV-YLPKPSM 155

RESULT 21

AATM_MOUSE STANDARD: PRT: 430 AA.

ID AATM_MOUSE

AC P05202; 009188; (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)

DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).

GN GOT2 OR GOT-2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=87057413; PubMed=3782150; Nagashima F., Morino Y., Obaru K., Nomiyama H., Shimada K., Setoyama C., Shimada K.;

RT "Cloning and sequence analysis of mRNA for mouse aspartate aminotransferase isoenzymes."

RT J. Biol. Chem. 261:16976-16983(1986).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=C3H/He; TISSUE=Liver;

RX MEDLINE=88118911; PubMed=2828632; Tsuzuki T., Obaru K., Setoyama C., Shimada K.;

RT "Structural organization of the mouse mitochondrial aspartate aminotransferase gene."

RT J. Mol. Biol. 198:21-31(1987).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=FVB; TISSUE=Liver;

RA Bradbury M.W., Berk P.D.;

RT Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.

CC -1 CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate + L-glutamate.

CC -1 COFACTOR: PYRIDOXAL PHOSPHATE.

CC -1 SUBUNIT: HOMODIMER.

CC -1 SUBCELLULAR LOCATION: Mitochondrial matrix.

CC -1 MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC ONE AND A MITOCHONDRIAL ONE.

CC -1 SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT AMINOTRANSFERASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: J02622; AAA37264.1; -

DR EMBL: X06917; CAA30015.1; -

DR EMBL: X06918; CAA30015.1; JOINED.

DR EMBL: X06919; CAA30015.1; JOINED.

DR EMBL: X06920; CAA30015.1; JOINED.

DR EMBL: X06921; CAA30015.1; JOINED.

DR EMBL: X06922; CAA30015.1; JOINED.

DR EMBL: X06923; CAA30015.1; JOINED.

DR EMBL: X06924; CAA30015.1; JOINED.

DR EMBL: X06925; CAA30015.1; JOINED.

DR EMBL: X06926; CAA30015.1; JOINED.

DR EMBL: M37259; AAA37265.1; ALT-SEQ.

DR EMBL: M37250; AAA37265.1; JOINED.

DR EMBL: M37251; AAA37265.1; JOINED.

DR EMBL: M37252; AAA37265.1; JOINED.

DR EMBL: M37253; AAA37265.1; JOINED.

DR EMBL: M37254; AAA37265.1; JOINED.

DR EMBL: M37255; AAA37265.1; JOINED.

DR EMBL: M37256; AAA37265.1; JOINED.

DR EMBL: M37258; AAA37265.1; JOINED.

DR EMBL: U82470; AAB91426.1; -

DR PIR: S01174; S01174.

DR PIR: A25349; A25349.

DR HSSP: P00508; TAAT.

DR MGD: MGI:95792; GOT2.

DR InterPro: IPR004839; AminoTransf1/2.

DR InterPro: IPR000796; Aspartate sub.

DR InterPro: IPR004838; NHTransf_1.

DR Pfam: PF00155; aminotran_1_2; 1.

DR PRINTS: PRO0199; TRANSAMINASE.

DR PROSITE: PS00105; AA-TRANSFER-CLASS-1; 1.

DR TRANSFASER: Aminotransferase; Pyridoxal phosphate; Mitochondrion; Transf peptide.

FW TRANSIT 1 29

FT CHAIN 30 430

FT BINDING 279 279

FT CONFLICT 146 146

FT SEQUENCE 430 AA; 47411 MW; D590524CA/FFB885 CRC64;

Query Match 27.4%; Score 45.5; DB 1; Length 430;

Best Local Similarity 44.0%; Pred. No. 38;

Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;

OY 3 RSGSEVQGLFGHPDVFYLMKPM 27

DB 139 RVGASFQRFKFSRDV-FLPKPSM 162

RESULT 22

AATM_RAT STANDARD: PRT: 430 AA.

ID AATM_RAT

AC P00507; Q64551; (Rel. 01, Created)

DT 21-JUL-1986 (Rel. 01, Last sequence update)

DT 01-JUL-1989 (Rel. 11, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)

DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).

GN GOT2 OR MAAT.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=88106546; PubMed=3322287; Martinez-Carrion M.;

RA Martinez-Carrion M.;

RT "Molecular cloning and in vivo expression of a precursor to rat mitochondrial aspartate aminotransferase."

RT Biochem. Biophys. Res. Commun. 149:859-865(1987).

RN [2]

RP SEQUENCE OF 30-430.

RX MEDLINE=81133608; PubMed=7470110; Hyunh Q.K., Sakakibara R., Watanabe T., Wada H.;

RA Hyunh Q.K., Sakakibara R., Watanabe T., Wada H.;

RT "Primary structure of mitochondrial glutamic oxaloacetic transaminase from rat liver: comparison with that of the pig heart isozyme."

RT Biochem. Biophys. Res. Commun. 97:474-479(1980).

RN [3]

RP SEQUENCE OF 30-430.

RX MEDLINE=82075716; PubMed=7309704; Hyunh Q.K., Sakakibara R., Watanabe T., Wada H.;

RA Hyunh Q.K., Sakakibara R., Watanabe T., Wada H.;

RT "The complete amino acid sequence of mitochondrial glutamic oxaloacetic transaminase from rat liver."

RT J. Biochem. 90:863-875(1981).

RN [4]

RP SEQUENCE OF 1-30 FROM N.A.

RC STRAIN=Wistar; TISSUE=Liver;

RX MEDLINE=95279397; PubMed=7759512; Juang H.H., Costello L.C., Franklin R.B.;

RA Juang H.H., Costello L.C., Franklin R.B.;

RT "Androgen modulation of multiple transcription start sites of the mitochondrial aspartate aminotransferase gene in rat prostate."


```

RN J. Biol. Chem. 270:12629-12634(1995).
RN [5]
RP SEQUENCE OF 279-329 FROM N.A.
RX MEDLINE-86130567; PubMed-3004464;
RA Horio Y., Sakakibara R., Tanaka T., Taketoshi M., Obaru K.,
RA Shimada K., Morino Y., Wada H.;
RT "Molecular cloning of rat mitochondrial glutamic oxaloacetic
RT transaminase mRNA and regulation of its expression in regenerating
RT liver."
RL Biochem. Biophys. Res. Commun. 134:803-811(1986).
CC -1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
CC L-glutamate.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
CC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC AMINOTRANSFERASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M18467; AAB54275.1; -
DR EMBL: U21158; AAC13868.1; -
DR EMBL: M12709; AAA41267.1; -
DR PIR: A28005; XNRDM.
DR HSSP: P00508; 7AAT.
DR InterPro: IPR004839; AminoTransfl/2.
DR InterPro: IPR000796; AsptTransf_1.
DR InterPro: IPR004838; NHTransf_1.
DR Pfam: PF00155; aminotran.1.2; 1.
DR PRINTS: PR00799; TRANSAMINASE.
DR PROSITE: PS00105; AL-TRANSFER_CLASS_1; 1.
KW Transferrase; AminoTransferase; Pyridoxal phosphate; Mitochondrion;
KW Transfl peptide.
FT TRANSIT 1 29
FT CHAIN 30 430
FT BINDING 279 279
FT CONFLICT 30 30
FT CONFLICT 162 162
FT CONFLICT 167 167
FT CONFLICT 177 177
FT CONFLICT 232 232
FT CONFLICT 255 255
FT CONFLICT 338 339
FT CONFLICT 352 352
FT CONFLICT 386 386
FT CONFLICT 400 400
SQ SEQUENCE 430 AA; 47314 MW; EDCB862A20DB736 CRC64;
Query Match 27.4%; Score 45.5; DB 1; Length 430;
Best Local Similarity 44.0%; Pred. No. 38;
Matches 11; Conservative 5; Mismatches 8; Indels 1; Gaps 1;
OY 3 RSGSSFGVGLFGQHPDYFYMEPAW 27
Db 139 RVGASFQRFKFSRDVF-LPKPSW 162

```

```

GN KUP OR TRKD OR B3747.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE-94042856; PubMed-8226635;
RA Schleyer M., Bakker E.P.;
RT "Nucleotide sequence and 3'-end deletion studies indicate that the
RT K(+) uptake protein kup from Escherichia coli is composed of a
RT hydrophobic core linked to a large and partially essential
RT hydrophilic C terminus."
RL J. Bacteriol. 175:6925-6931(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE-93315143; PubMed-7686882;
RA Burland V.D., Plunkett G. III, Daniels D.L., Blattner F.R.;
RT "DNA sequence and analysis of 136 kilobases of the Escherichia coli
RT genome: organizational symmetry around the origin of replication."
RT [Genomics 15:551-561(1993)].
RN REVISIONS.
RP STRAIN-K12 / MG1655;
RX MEDLINE-94089392; PubMed-8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
CC -1- FUNCTION: LOW-AFFINITY POTASSIUM TRANSPORT SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: TO S.OCCIDENTALIS POTASSIUM TRANSPORT HAK1.
CC -1- CAUTION: REF.3 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 506.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68551; AAA48555.1; -
DR EMBL: L10328; AAA62100.1; ALT_FRAME.
DR EMBL: AE000451; AAC76770.1; ALT_FRAME.
DR PIR: S31717; S31717.
DR PIR: A49345; A49345.
DR Ecogene; ES11541; Kup.
DR InterPro: IPR003855; K+-transprt.
DR Pfam: PF02705; K_transf.1.
DR TIGRfam; TIGR00794; kup; 1.
KW Transmembrane; Inner membrane; Transport; Potassium transport;
KW Complete proteome.
FT DOMAIN 1 8
FT TRANSMEM 9 29
FT DOMAIN 30 47
FT TRANSMEM 48 66
FT DOMAIN 67 101
FT TRANSMEM 102 123
FT DOMAIN 124 133
FT TRANSMEM 134 157
FT DOMAIN 158 167
FT TRANSMEM 168 185
FT DOMAIN 186 211
FT TRANSMEM 212 232
FT DOMAIN 233 244
FT TRANSMEM 245 267
FT DOMAIN 268 283
FT TRANSMEM 284 304

```

```

FT DOMAIN 305 336 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 337 357 POTENTIAL.
FT DOMAIN 358 362 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 363 383 POTENTIAL.
FT TRANSMEM 384 394 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 395 414 POTENTIAL.
FT DOMAIN 415 418 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 419 440 POTENTIAL.
FT DOMAIN 441 622 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 622 AA; E556791ABF98BD5 CRC64;

```

Query Match
Best Local Similarity 27.4%; Score 45.5; DB 1; Length 622;
Matches 10; Conservative 8; Mismatches 9; Indels 19; Gaps 1;

```

QY 3 RSGSFTVGQLFGF-----QHPDVFLYMEPAWH 29
DB 159 KGTANVGKLFAPIMTWELLAGLGRSTIANPEVLAHNPMAV 204

```

RESULT 24

```

ID TRAB_HUMAN STANDARD; PRT; 708 AA.
AC Q9UGI0;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE TRABID protein.
GN TRABID.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Evans P.C., Coagwell W.J., Taylor E., Kilshaw P.J.;
RT "Modulation of TNF alpha signalling by a novel protein TRABID."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC 1- FUNCTION: MODULATES TNF ALPHA SIGNALING.
CC 1- SIMILARITY: CONTAINS 3 RANBP2-TYPE ZINC FINGERS.
CC 1- SIMILARITY: CONTAINS 1 OTU DOMAIN.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AJ252060; CAB64449.1; -
CC InterPro: IPR003323; OTU.
CC InterPro: IPR001876; Znf_RanGDP.
CC Pfam: PF00641; Zf-RanBP_2.
CC SMART: SM00547; Znf_RB2_3.
CC PROSITE: PS50602; OTU_1.
CC PROSITE: PS01358; ZF-RANBP2_1; 3.
CC PROSITE: PS50199; ZF-RANBP2_2; 3.
CC Zinc-finger; Repeat.
CC ZN_FING 3 33 RANBP2-TYPE 1.
CC ZN_FING 84 113 RANBP2-TYPE 2.
CC ZN_FING 149 178 RANBP2-TYPE 3.
CC DOMAIN 432 592 OTU
CC SEQUENCE 708 AA; 80906 MW; DE6F7DADD2DE23A CRC64;

```

Query Match
Best Local Similarity 27.4%; Score 45.5; DB 1; Length 708;
Matches 10; Conservative 2; Mismatches 11; Indels 3; Gaps 1;

```

QY 2 WRSGSFTVGQLFGQHPDVFLYMEPAW 27
DB 481 WKDWSWYSQSGDL--FSLREQW 503

```

RESULT 25

```

ID CBR1_RHIME STANDARD; PRT; 313 AA.
AC P58332; Q9EXV3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Rubisco operon transcriptional regulator.
GN CBR OR RB0196 OR SBE20203.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxId=582;
RN [1]
RP SEQUENCE FROM N.A.
RA Finan T.M., Aneja P., Chain P., Napper K., Golding B.;
RT "Mineral phosphate solubilization in Sinorhizobium meliloti."
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.

```

```

CC 12)
CC SEQUENCE FROM N.A.
CC STRAIN-1021;
CC MEDLINE=21396508; PubMed=11481431;
CC Finan T.M., Weidner S., Wong K., Buhmester J., Chain P.,
CC Vorholter F.J., Hernandez-Lucas I., Becker A., Gouzy J.,
CC Golding B., Puehler A.;
CC "The complete sequence of the 1.683-kb psymb megaplasmid from the N2-
CC fixing endosymbiont Sinorhizobium meliloti."
CC Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
CC 1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR FOR THE CBG OPERON FOR
CC RUBISCO AND OTHER CALVIN CYCLE GENES (BY similarity).
CC 1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.

```

```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AY013584; AAC42537.1; -
CC EMBL: AL603642; CAC48596.1; -
CC InterPro: IPR000847; HTH_LysR.
CC InterPro: IPR005119; LysR_subst.
CC Pfam: PF00126; HTH_1; 1.
CC Pfam: PF03466; LysR_substrate; 1.
CC PRINTS: PR00039; HTHLYSR.
CC PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
CC Transcription regulation; Activator; DNA-binding; Plasmid;
CC KW Complete proteome.
CC DNA_BIND 21 40 H-T-H MOTIF (POTENTIAL).
CC SEQUENCE 313 AA; 34363 MW; 1BB07B3B46B829C7 CRC64;

```

Query Match
Best Local Similarity 27.1%; Score 45; DB 1; Length 313;
Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

```

QY 11 QLFQHPDVFLYMEPAWH 28
DB 160 CVFGDHPVLF--IAPAGH 175

```

RESULT 26

```

ID CBR2_RHIME STANDARD; PRT; 313 AA.
AC P56885;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```

```

DE Rubisco operon transcriptional regulator.
GN CBR.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
CC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSM419;
RA Penner B.J., Tiwari R.P., Dilworth M.J.;
RT "Genetic regulation of C2 metabolism in Sinorhizobium meliloti.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: TRANSCRIPTIONAL ACTIVATOR FOR THE CBB OPERON FOR
CC RUBISCO AND OTHER CALVIN CYCLE GENES.
CC -1- SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC REGULATORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF211846; AAF25374.1; -
CC InterPro: IPR000847; HTH_LysR.
CC InterPro: IPR005119; LysR_subst.
CC Pfam: PF00126; HTH.1; 1.
CC Pfam: PF03466; LysR_substrate; 1.
CC PRINTS: PR00039; HHLYSR.
CC PROSITE: PS00044; HTH_LYSR_FAMILY; 1.
CC Transcription regulation; Activator; DNA-binding.
KW DNA_BIND 21 40 H-T-H MOTIF (POTENTIAL).
FT SQUENCE 313 AA; 34192 MW; 818CF5EFC0CED24 CRC64;
SQ
Query Match 27.1%; Score 45; DB 1; Length 313;
Best Local Similarity 55.6%; Pred. No. 32;
Matches 10; Conservative 2; Mismatches 4; Indels 2; Gaps 1;
OY 11 QLFQGHDPVYLYMEPAWH 28
DB 160 QVFQDHLVF-IAPAGH 175

```

```

RESULT 27
AROC_STYX3 STANDARD; PRT; 362 AA.
AC P23353;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chorismate synthase (PC 4.6.1.4) (5-enolpyruvylshikimate-3-phosphate
DE phospholase).
GN AROC OR SL11747.
OS Synecocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=94086566; PubMed=7505271;
RA Schmidt J., Bubunenko M., Subramanian A.R.;
RT "A novel operon organization involving the genes for chorismate
RT synthase (aromatic biosynthesis pathway) and ribosomal GTPase center
RT proteins (L11, L1, L10, L12: rplKarl) in cyanobacterium Synecocystis
RT PCC 6803.";
RL J. Biol. Chem. 268:27447-27457(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=97061201; PubMed=8905231;
RA Keneo T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Natsu K.,

```

```

RA Okumura S., Shimo S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
RN [3]
RP SEQUENCE OF 1-138 FROM N.A.
RC MEDLINE=91002677; PubMed=2119815;
RA Sibold C., Subramanian A.R.;
RT Cloning and characterization of the genes for ribosomal proteins L10
RT and L12 from Synecocystis sp. PCC 6803: comparison of gene
RT clustering pattern and protein sequence homology between
RT cyanobacteria and chloroplasts.";
RL Biochim. Biophys. Acta 1050:61-68(1990).
CC -1- CATALYTIC ACTIVITY: 5-O-(1-carboxyvinyl)-3-phosphoshikimate =
CC chorismate + phosphate.
CC -1- COFACTOR: REDUCED FLAVIN (BY SIMILARITY).
CC -1- PATHWAY: Aromatic amino acids biosynthesis; shikimate pathway;
CC seventh step.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CHORISMATE SYNTHASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: X67516; CAA47855.1; -
CC EMBL: D90906; BAA17415.1; -
CC EMBL: X53178; CAA37319.1; -
CC PIR: S13070; S13070.
CC InterPro: IPR000453; Chorismate-synt.
CC Pfam: PF01264; Chorismate-synt.
CC PRODOM: PD002941; Chorismate-synt; 1.
CC TIGRPFAM: TIGR00033; aroc; 1.
CC PROSITE: PS00787; CHORISMATE_SYNTHASE_1; 1.
CC PROSITE: PS00788; CHORISMATE_SYNTHASE_2; 1.
CC PROSITE: PS00789; CHORISMATE_SYNTHASE_3; 1.
KW Lyase; Aromatic amino acid biosynthesis; Complete proteome.
SQ SEQUENCE 362 AA; 39287 MW; 9709B0F409168AB8 CRC64;
Query Match 27.1%; Score 45; DB 1; Length 362;
Best Local Similarity 50.0%; Pred. No. 37;
Matches 11; Conservative 1; Mismatches 8; Indels 2; Gaps 1;
OY 5 GSGFVGOLF--GOHPVFTIME 24
DB 249 GSGFAGTLLTGSQHNDEYILDE 270

```

```

RESULT 28
OPSL_SCHGR STANDARD; PRT; 381 AA.
AC Q94741;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Opsln 1.
GN IOL.
OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Psterygota; Neoptera; Orthoptera; Orthoptera; Celliara;
OC Acridoidea; Acrididae; Acrididae; Cyrtacanthacridinae;
OC Schistocerca.
OX NCBI_TaxID=7010;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97301174; PubMed=9156194;
RA Townner P., Harris P., Wolstenholme A.J., Hill C., Worm K., Gartner W.;

```

RT "Primary structure of locust opsins: a speculative model which may
 account for ultraviolet wavelength light detection.";
 RT Vision Res. 37:495-503(1997).
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY
 CC LINKED TO CIS-RETINAL.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY
 CC BE PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC OPSIN SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X80071; CAA56377.1; -
 DR HSSP: P02699; 1F88.
 DR InterPro: IPR000276; GPCR_Rhodopsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCR_Rhodopsn.
 DR PROSITE: PS00237; G-PROTEIN_RECPT_FL1; 1.
 DR PROSITE: PS50262; G-PROTEIN_RECPT_FL2; 1.
 DR PROSITE: PS00238; OPSIN; 1.
 KW Photoreceptor; Retinal protein; Transmembrane; Phosphorylation;
 KW Glycoprotein; G-protein coupled receptor; Vision.
 FT DOMAIN 1 53
 FT TRANSMEM 54 78 1 (POTENTIAL).
 FT DOMAIN 79 90 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 91 115 2 (POTENTIAL).
 FT DOMAIN 116 130 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 131 150 3 (POTENTIAL).
 FT DOMAIN 151 169 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 170 193 4 (POTENTIAL).
 FT DOMAIN 194 217 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 218 245 5 (POTENTIAL).
 FT DOMAIN 246 280 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 281 304 6 (POTENTIAL).
 FT DOMAIN 305 311 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 312 336 7 (POTENTIAL).
 FT DOMAIN 337 381 CYTOPLASMIC (POTENTIAL).
 FT DISULFID 127 204 BY SIMILARITY.
 FT BINDING 323 323 RETINAL CHROMOPHORE (BY SIMILARITY).
 FT CARBOHYD 24 24 N-LINKED (GLCNAC...) (POTENTIAL).
 FT CARBOHYD 200 200 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 381 AA; 42398 MW; 50D38FF9038D5BA5 CRC64;
 Query Match 27.1%; Score 45; DB 1; Length 381;
 Best Local Similarity 27.6%; Pred. No. 39;
 Matches 8; Conservative 5; Mismatches 14; Indels 2; Gaps 1;
 QY 2 WRSGSFVGLFGQH--PDVFILMEPAWH 28
 DB 16 WGSAGFANQTVVDPVPEMLYLPDHW 44
 RESULT 29
 FIBA_HUMAN STANDARD: PRT; 866 AA.
 AC P02671; Q9BX62; Q9UCH2;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Fibrinogen alpha/alpha-E chain precursor [contains: Fibrinopeptide A].
 GN FGA.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ALPHA-E FORM).
 RX MEDLINE=93090725; PubMed=1457396;
 RA Fu Y., Weissbach L., Plant P.W., Oddoux C., Cao Y., Liang T.J.,
 RA Roy S.N., Reiman C.M., Grieninger G.;
 RT "Carboxy-terminal extended variant of the human fibrinogen alpha
 RT subunit: a novel exon conferring marked homology to beta and gamma
 RT subunits.";
 RL Biochemistry 31:11968-11972(1992).
 RN [2]
 RP SEQUENCE FROM N.A. (ALPHA-E FORM).
 RA Chung D.W., Grieninger G.;
 RT "Fibrinogen DNA and protein sequences.";
 RL (in) Ebert R.F. (eds.);
 RL Index of variant human fibrinogens, pp.13-24, CRC Press,
 RL Boca Raton (1994).
 RN [3]
 RP SEQUENCE FROM N.A. (ALPHA-E; ALPHA), AND VARIANTS VAL-6; ALA-331 AND
 RP ALA-456.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1-655 FROM N.A. (ALPHA-E FORM).
 RX TISSUE=Liver;
 RT MEDLINE=91344740; PubMed=2102623;
 RA Chung D.W., Harris J.E., Davie E.W.;
 RT "Nucleotide sequences of the three genes coding for human
 RT fibrinogen.";
 RL Adv. Exp. Med. Biol. 281:39-48(1990).
 RN [5]
 RP SEQUENCE FROM N.A. (ALPHA FORM).
 RX MEDLINE=83247396; PubMed=6575389;
 RA Kant J.A., Lord S.T., Crabtree G.R.;
 RT "Partial mRNA sequences for human A alpha, B beta, and gamma
 RT fibrinogen chains: evolutionary and functional implications.";
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3953-3957(1983).
 RN [6]
 RP SEQUENCE OF 1-629 FROM N.A.
 RX MEDLINE=83268432; PubMed=6688355;
 RA Rixon M.W., Chan W.-Y., Davie E.W., Chung D.W.;
 RT "Characterization of a complementary deoxyribonucleic acid coding for
 RT the alpha chain of human fibrinogen.";
 RL Biochemistry 22:3237-3244(1983).
 RN [7]
 RP SEQUENCE OF 20-629.
 RA Henschen A., Iottspeich F., Southan C., Topfer-Petersen E.;
 RT "Human fibrinogen: sequence, sulfur bridges, glycosylation and some
 RT structural variants.";
 RL (in) Peeters H. (eds.);
 RL Protides of the biological fluids, Proc. 28th colloquium, pp.51-56,
 RL Pergamon Press, Oxford (1980).
 RN [8]
 RP SEQUENCE OF 20-629, AND DISULFIDE BONDS.
 RX MEDLINE=80088231; PubMed=518846;
 RA Walt K.W.K., Cottrell B.A., Strong D.D., Doonittle R.F.;
 RT "Amino acid sequence studies on the alpha chain of human
 RT fibrinogen: overlapping sequences providing the complete sequence.";
 RL Biochemistry 18:5410-5416(1979).
 RN [9]
 RP SEQUENCE OF 110-156 FROM N.A.
 RX MEDLINE=84069777; PubMed=6689067;
 RA Imam A.M., Eaton M.A., Williamson R., Humphries S.;
 RT "Isolation and characterisation of cDNA clones for the A alpha- and
 RT gamma-chains of human fibrinogen.";
 RL Nucleic Acids Res. 11:7427-7434(1983).
 RN [10]
 RP SEQUENCE OF 605-644 FROM N.A. (ALPHA FORM).
 RX MEDLINE=83254384; PubMed=6575700;
 RA Chung D.W., Rixon M.W., Que B.G., Davie E.W.;
 RT "Cloning of fibrinogen genes and their cDNA.";
 RL Ann. N.Y. Acad. Sci. 408:449-456(1983).

RN [11] SEQUENCE OF 20-35.
RA Blomback B., Blomback M., Gron Dahl N.U., Guthrie C., Hinton M.;
RT "Studies of fibrinopeptides from primates.";
RL Acta Chem. Scand. 19:1788-1789(1965).
RN [12] CROSS-LINKING ACCEPTOR SITES.
RX MEDLINE=80088230; PubMed=518845;
RA Cottrell B.A., Strong D.D., Watt K.W.K., Doolittle R.F.;
RT "Amino acid sequence studies on the alpha chain of human fibrinogen.
RT Exact location of cross-linking acceptor sites.";
RL Biochemistry 18:5405-5410(1979).
RN [13] CROSS-LINKING ACCEPTOR SITES.
RX MEDLINE=78130085; PubMed=632262;
RA Fretto L.J., Ferguson E.W., Steinman H.M., McKee P.A.;
RT "Localization of the alpha-chain cross-link acceptor sites of human
RT fibrin.";
RL J. Biol. Chem. 253:2184-2195(1978).
RN [14] VARIANT, AND DISULFIDE BONDS.
RX MEDLINE=76225080; PubMed=936108;
RA Blomback B., Hessel B., Hogg D.;
RT "Disulfide bridges in NH2-terminal part of human fibrinogen.";
RL Thromb. Res. 8:639-658(1976).
RN [15] REVISED, EM STRUCTURE, POLYMERIZATION, AND LIGANDS.
RX MEDLINE=84305751; PubMed=6383194;
RA Doolittle R.F.;
RT "Fibrinogen and fibrin.";
RL Annu. Rev. Biochem. 53:195-229(1984).
RN [16] CROSS-LINKING SITE FOR ALPHA-2-PLASMIN INHIBITOR.
RX MEDLINE=87057190; PubMed=2877981;
RA Kimura S., Aoki N.;
RT "Cross-linking site in fibrinogen for alpha 2-plasmin inhibitor.";
RL J. Biol. Chem. 261:15591-15595(1986).
RN [17] PHOSPHORYLATION.
RX MEDLINE=84104274; PubMed=6318767;
RA Itarte E., Plana M., Guasch M.D., Martos C.;
RT "Phosphorylation of fibrinogen by casein kinase I.";
RL Biochem. Biophys. Res. Commun. 117:631-636(1983).
RN [18] X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-39.
RX MEDLINE=92218459; PubMed=1560020;
RA Martin P.D., Robertson W., Turk D., Huber R., Bode W., Edwards B.F.P.;
RT "The structure of residues 7-16 of the A alpha-chain of human
RT fibrinogen bound to bovine thrombin at 2.3-A resolution.";
RL J. Biol. Chem. 267:7911-7920(1992).
RN [19] X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 130-216.
RX MEDLINE=97472408; PubMed=9333233;
RA Spraggon G., Everse S.J., Doolittle R.F.;
RT "Crystal structures of fragment D from human fibrinogen and its
RT crosslinked counterpart from fibrin.";
RL Nature 389:455-462(1997).
RN [20] X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 130-216.
RX MEDLINE=98292395; PubMed=9628725;
RA Everse S.J., Spraggon G., Veerapandian L., Riley M., Doolittle R.F.;
RT "Crystal structure of fragment double-D from human fibrin with two
RT different bound ligands.";
RL Biochemistry 37:8637-8642(1998).
RN [21] X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 670-866.
RX MEDLINE=98356117; PubMed=9689040;
RA Spraggon G., Applegate D., Everse S.J., Zhang J.Z., Veerapandian L.,
RA Redman C., Doolittle R.F., Griening G.;
RT "Crystal structure of a recombinant alpha2C domain from human
RT fibrinogen-420.";
RL Proc. Natl. Acad. Sci. U.S.A. 95:9099-9104(1998).
RN [22]

RP X-RAY CRYSTALLOGRAPHY.
RX MEDLINE=99175089; PubMed=10074346;
RA Everse S.J., Spraggon G., Veerapandian L., Doolittle R.F.;
RT "Conformational changes in fragments D and double-D from human
RT fibrin(ogen) upon binding the peptide ligand Gly-His-Pro-amide.";
RL Biochemistry 38:2941-2946(1999).
RN [23] VARIANT KYOTO-2.
RX MEDLINE=91300048; PubMed=2070049;
RA Yoshida N., Okuma M., Hirata H., Matsuda M., Yamazumi K., Asakura S.;
RT "Fibrinogen Kyoto II, a new congenitally abnormal molecule,
RT characterized by the replacement of A alpha proline-18 by leucine.";
RL Blood 78:149-153(1991).
RN [24] VARIANT LIMA.
RX MEDLINE=92340680; PubMed=1634621;
RA Maekawa H., Arocha-Pinango C.L., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., Arocha-Pinango C.L., Rodriguez S., Nagy H.,
RA Perez-Requejo J.L., Matsuda M.;
RT "Fibrinogen Lima: a homozygous dysfibrinogen with an A
RT alpha-arginine-141 to serine substitution associated with extra
RT N-glycosylation at A alpha-asparagine-139. Impaired fibrin gel
RT formation but normal fibrin-facilitated plasminogen activation
RT catalyzed by tissue-type plasminogen activator.";
RL J. Clin. Invest. 90:67-76(1992).
RN [25] VARIANT CARACAS-2.
RX MEDLINE=91268018; PubMed=1675636;
RA Maekawa H., Yamazumi K., Muramatsu S., Kaneko M., Hirata H.,
RA Takahashi N., de Bosch N.B., Carvajal Z., Ojeda A.,
RA Arocha-Pinango C.L., Matsuda M.;
RT "An A alpha Ser-434 to N-glycosylated Asn substitution in a
RT dysfibrinogen, fibrinogen Caracas II, characterized by impaired
RT fibrin gel formation.";
RL J. Biol. Chem. 266:11575-11581(1991).
RN [26] VARIANT DUSART.
RX MEDLINE=93323289; PubMed=8473507;
RA Koopman J., Haverkate F., Grimbergen J., Lord S.T., Mosseson M.W.,
RA Diorio J.P., Stebenlist K.S., Legrand C., Soria J., Soria C.,
RA Caen J.P.;
RT "Molecular basis for fibrinogen Dusart (A alpha 554 Arg-->Cys) and
RT its association with abnormal fibrin polymerization and
RT thrombophilia.";
RN [27] Query Match 27.1%; Score 45; DB 1; Length 866;
Best Local Similarity 52.9%; Pred. No. 93;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;
QY 1 SWRSGSFVQILGQHP 17
DB 333 SWRSGSGTGTGNGNP 349
RESULT 30
CHRD_BRARE STANDARD; PRT; 940 AA.
AC 057472; OSDEDB;
DT 16-OCT-2001 (rel. 40, Created)
DT 15-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Chordin precursor (Chordin protein).
GN CHD.
OS Brachydanio rerio (zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxId=7953;
RN [1]
RP SEQUENCE FROM N.A., FUNCTION, TISSUE SPECIFICITY, AND DEVELOPMENTAL
RP STAGE.
RC TISSUE-gastrula;
RX MEDLINE=98104254; PubMed=9441687;

RA Miller-Bertoglio V.E., Fisher S., Sanchez A., Mullins M.C.,
RA Halpern M.E.;
RT "Differential regulation of chordin expression domains in mutant
zebrafish";
RL Dev. Biol. 192:537-550(1997).
RN
RP
SEQUENCE OF 1-42 FROM N.A.
RA Fujii R., Hibi M., Hirano T., Shimizu T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Dorsalizing factor. Key developmental protein that
CC dorsalizes early vertebrate embryonic tissues by binding to
CC ventralizing TGF-beta family bone morphogenetic proteins (BMPs)
CC and sequestering them in latent complexes (By similarity).
CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
CC -1- TISSUE SPECIFICITY: During gastrulation, levels are highest in
CC the organizer region. Also present in the developing brain and in
CC paraxial mesoderm and ectoderm.
CC -1- DEVELOPMENTAL STAGE: First detected shortly after the midblastula
CC transition. Levels increase during gastrulation, persist through
CC early somitogenesis, but then decrease and are gone by 24
CC hours.
CC -1- SIMILARITY: BELONGS TO THE CHORDIN FAMILY.
CC -1- SIMILARITY: CONTAINS 4 WMFC DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: AF034606; AAB93485.1; -;
CC DR EMBL: AB043968; BAB18643.1; -;
CC DR ZFIN: ZDB-GENE-990415-33; chd.
CC DR InterPro: IPR001007; WMF_C.
CC DR Pfam: PF00093; WWC; 4.
CC DR SMART: SM00214; WWC; 4.
CC DR PROSITE: PS01208; WWC; 3.
CC DR Developmental protein: Repeat: Glycoprotein: Signal.
CC FT SIGNAL 1 19 POTENTIAL.
CC FT CHAIN 20 940 CHORDIN.
CC FT DOMAIN 42 118 WMFC 1.
CC FT DOMAIN 689 748 WMFC 2.
CC FT DOMAIN 767 836 WMFC 3.
CC FT DOMAIN 855 919 WMFC 4.
CC FT CARBOHYD 347 347 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT CARBOHYD 430 430 N-LINKED (GLCNAC...) (POTENTIAL).
CC FT SEQUENCE 940 AA, 104999 MW, B855CAF8F4F623AC CRC64;
SQ
Query Match 27.1%; Score 45; DB 1; Length 940;
Best Local Similarity 40.0%; Pred. No. 1e+02;
Matches 10; Conservative 2; Mismatches 5; Indels 8; Gaps 1;
QY 4 SSSPFGOLFQHPDYFYIMEPAWH 28
DB 42 SSGSFGGR-----FYLEDITWH 58
RESULT 31
PRED HUMAN STANDARD; PRT; 1912 AA.
ID P23468;
AC P23468;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Protein-tyrosine phosphatase delta precursor (BC 3.1.3.48) (R-PTP-
DE delta).
GN PTPRD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A., AND MUTAGENESIS OF ARG-1178.
RX MEDLINE=95204468; PubMed=7896816;
RA Pulido R., Krueger N.X., Serra-Pages C., Saito H., Streuli M.;
RT "Molecular characterization of the human transmembrane protein -
RT tyrosine phosphatase delta. Evidence for tissue-specific expression of
RT alternative human transmembrane protein-tyrosine phosphatase delta
RT isoforms";
RL J. Biol. Chem. 270:6722-6728(1995).
RN [2]
RP SEQUENCE OF 390-1912 FROM N.A.
RX TISSUE=Placenta;
RA MEDLINE=91060018; PubMed=2170109;
RT Krueger N.X., Streuli M., Saito H.;
RT "Structural diversity and evolution of human receptor-like protein
RT tyrosine phosphatases";
RL EMBO J. 9:3241-3252(1990).
CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphate + H(2)O = protein
CC tyrosine + phosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS: A number of isoforms are produced by
CC alternative splicing.
CC -1- PTM: A CLEAVAGE OCCURS THAT SEPARATES THE EXTRACELLULAR DOMAIN
CC FROM THE TRANSMEMBRANE SEGMENT.
CC -1- SIMILARITY: CONTAINS 3 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC -1- SIMILARITY: CONTAINS 8 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 PROTEIN-TYROSINE PHOSPHATASE DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC DR EMBL: L38929; AAC41749.1; -;
CC DR EMBL: X54133; CAA38068.1; -;
CC DR PIR: S12052; S12052.
CC DR HSSP: P18052; IYFO.
CC DR GeneW: HGNC:9668; PTPRD.
CC DR MIM: 601598; -;
CC DR InterPro: IPR003961; FN_III.
CC DR InterPro: IPR003962; FNIII_repeat.
CC DR InterPro: IPR003006; Ig_MHC.
CC DR InterPro: IPR003598; Ig_C2.
CC DR InterPro: IPR000387; TYR_phosphatase.
CC DR InterPro: IPR000242; TYR_PP.
CC DR Pfam: PF00041; fn3; 8.
CC DR Pfam: PF00047; fn3; 8.
CC DR Pfam: PF00102; Y_phosphatase; 2.
CC DR PRINTS: PR00014; FNTYPEIII.
CC DR PRINTS: PR00700; PRTYPEPHASE.
CC DR SMART: SM00060; FN3; 8.
CC DR SMART: SM00408; IGC2; 3.
CC DR SMART: SM0184; PTPC; 2.
CC DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 2.
CC DR PROSITE: PS50056; TYR_PHOSPHATASE_2; 2.
CC DR PROSITE: PS50055; TYR_PHOSPHATASE_PP; 2.
CC KW Hydroxylase; Receptor; Glycoprotein; Signal; Transmembrane; Repeat;
CC Immunoglobulin domain; Alternative splicing.
CC FT SIGNAL 1 20 POTENTIAL.
CC FT CHAIN 21 1912 PROTEIN-TYROSINE PHOSPHATASE DELTA.
CC FT DOMAIN 21 1265 EXTRACELLULAR (POTENTIAL).
CC FT TRANSSEM 1266 1290 POTENTIAL.
CC FT DOMAIN 1291 1912 CYTOPLASMIC (POTENTIAL).
CC FT DOMAIN 23 115 IG-LIKE C2-TYPE DOMAIN 1.
CC FT DOMAIN 118 225 IG-LIKE C2-TYPE DOMAIN 2.
CC FT DOMAIN 232 318 IG-LIKE C2-TYPE DOMAIN 3.
CC FT DOMAIN 320 414 FIBRONECTIN TYPE-III 1.
CC FT DOMAIN 417 513 FIBRONECTIN TYPE-III 2.
CC FT DOMAIN 516 606 FIBRONECTIN TYPE-III 3.
CC FT DOMAIN 609 708 FIBRONECTIN TYPE-III 4.

FT DOMAIN 711 822 FIBRONECTIN TYPE-III 5.
 FT DOMAIN 825 916 FIBRONECTIN TYPE-III 6.
 FT DOMAIN 918 1017 FIBRONECTIN TYPE-III 7.
 FT DOMAIN 1020 1137 FIBRONECTIN TYPE-III 8.
 FT DOMAIN 1375 1618 PROTEIN-TYROSINE PHOSPHATASE 1.
 FT DOMAIN 1619 1912 PROTEIN-TYROSINE PHOSPHATASE 2.
 FT ACT_SITE 1553 1553 BY SIMILARITY.
 FT ACT_SITE 1844 1844 BY SIMILARITY.
 FT SITE 1175 1178 CLEAVAGE (POTENTIAL).
 FT CARBOXYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXYD 299 299 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXYD 724 724 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOXYD 832 832 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 181 189 MISSING (IN KIDNEY ISOFORM).
 FT VARSPLIC 226 229 MISSING (IN KIDNEY ISOFORM).
 FT VARSPLIC 775 783 MISSING (IN KIDNEY ISOFORM).
 FT VARSPLIC 609 1137 MISSING (IN FETAL BRAIN ISOFORM).
 FT MTAGEN 1178 1178 R->A: 2.5-FOLD REDUCTION IN CLEAVAGE.
 SO SEQUENCE 1912 AA; 214759 MW; 3AE8C8D32182E26 CAC64;

Query Match 27.18; Score 45; DB 1; Length 1912;
 Best Local Similarity 35.7%; Pred. No. 2.1e+02;
 Matches 10; Conservative 4; Mismatches 12; Indels 2; Gaps 1;

QY 2 WRSSGSPVGFQGHDPVFLMEPAMHV 29
 DB 438 WKPEEPNGQIQYR--VYTMPTQHV 463

RESULT 32
 FAS_RAT STANDARD: PRT; 2505 AA.
 ID P12785; 064717; 009187; 009190;
 AC 01-OCT-1989 (Rel. 12, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DI 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Fatty acid synthase (EC 2.3.1.85) [includes: EC 2.3.1.39;
 EC 2.3.1.41; EC 1.1.1.100; EC 4.2.1.61; EC 1.3.1.10; EC 3.1.2.14].
 GN FASN.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89240686; PubMed=2717611;
 RA Amy C.M., Witkowski A., Naggert J., Williams B., Randhawa Z.,
 RA Smith S.,
 RT "Molecular cloning and sequencing of cDNAs encoding the entire rat
 RT fatty acid synthase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 86:3114-3118(1989).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Liver;
 RX MEDLINE=93075999; PubMed=3393331;
 RA Beck K.F., Schreglmann R., Stathopoulos I., Klein H., Hoch J.,
 RA Schweizer M.,
 RT "The fatty acid synthase (FAS) gene and its promoter in Rattus
 RT norvegicus.";
 RL DNA Seq. 2:359-366(1992).
 RN [3]
 RP SEQUENCE OF 75-2505 FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Mammary gland;
 RX MEDLINE=89128431; PubMed=2915923;
 RA Schweizer M., Takabayashi K., Beck K.F., Schreglmann R.,
 RT "Rat mammary gland fatty acid synthase: localization of the
 RT constituent domains and two functional polyadenylation/termination
 RT signals in the cDNA.";
 RL Nucleic Acids Res. 17:567-586(1989).
 RN [4]
 RP SEQUENCE OF 2085-2505 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=88087240; PubMed=2891707;

RA Naggert J., Witkowski A., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the thioesterase
 RT domain of the rat fatty acid synthetase.";
 RL J. Biol. Chem. 263:1146-1150(1988).
 RN [5]
 RP SEQUENCE OF 1921-2324 FROM N.A.
 RC TISSUE=Mammary gland;
 RX MEDLINE=87246646; PubMed=3109907;
 RA Witkowski A., Naggert J., Mikkelsen J., Smith S.;
 RT "Molecular cloning and sequencing of a cDNA encoding the acyl carrier
 RT protein and its flanking domains in the mammalian fatty acid
 RT synthetase.";
 RL Eur. J. Biochem. 165:601-606(1987).
 CC -1- FUNCTION: FATTY ACID SYNTHETASE CATALYZES THE FORMATION OF
 CC LONG-CHAIN FATTY ACIDS FROM ACETYL-COA, MALONYL-COA AND NADPH.
 CC THIS MULTIFUNCTIONAL PROTEIN HAS 7 CATALYTIC ACTIVITIES AND AN
 CC ACYL CARRIER PROPEIN.
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + N malonyl-CoA + 2N NADPH = a
 CC long-chain fatty acid + (N+1) CoA + N Co(2) + 2N NADP(+) = CoA +
 CC -1- CATALYTIC ACTIVITY: Acetyl-CoA + [acyl-carrier protein] = CoA +
 CC acetyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Malonyl-CoA + [acyl-carrier protein] = CoA +
 CC malonyl-[acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + malonyl-[acyl-
 CC carrier protein] = 3-oxoacyl-[acyl-carrier protein] + Co(2) +
 CC [acyl-carrier protein].
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxyacyl-[acyl-carrier protein] +
 CC NADP(+) = 3-oxoacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: (3R)-3-hydroxy-palmitoyl-[acyl-carrier protein]
 CC = 2-hexadecenoyl-[acyl-carrier protein] + H(2)O.
 CC -1- CATALYTIC ACTIVITY: Acyl-[acyl-carrier protein] + NADP(+) = trans-
 CC 2,3-dehydroacyl-[acyl-carrier protein] + NADPH.
 CC -1- CATALYTIC ACTIVITY: Oleoyl-[acyl-carrier protein] + H(2)O = [acyl-
 CC carrier protein] + oleate.
 CC -1- SUBUNIT: HOMODIMER, WHICH IS ARRANGED IN A HEAD TO TAIL FASHION.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-stb.ch/announce/>
 CC or send an email to license@isb-stb.ch).
 CC
 CC EMBL: M76767; AAA57219.1; -
 CC EMBL: X62888; CAA44679.1; -
 CC EMBL: X62889; CAA44680.1; -
 CC EMBL: X13415; CAA31780.1; -
 CC EMBL: X13527; CAA31882.1; -
 CC EMBL: J03514; AAA41144.1; -
 CC PIR: A30313; XYRTPA.
 DR InterPro: IPR001227; AC transferase.
 DR InterPro: IPR002085; Adh_zn_family.
 DR InterPro: IPR000784; ketoacyl_synth.
 DR InterPro: IPR003880; pantoic_acid.
 DR InterPro: IPR000051; SAM_bind.
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00107; adh_zinc.1.
 DR Pfam: PF00109; ketoacyl_synth.1.
 DR Pfam: PF00580; pp-binding.1.
 DR Pfam: PF00698; Acyl-transf.1.
 DR Pfam: PF00975; Thioesterase.1.
 DR PROSITE: PS02801; ketoacyl_synth_C.1.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE.1.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE.1.
 DR PROSITE: PS50075; ACP_DOMAIN.1.
 KM Fatty acid biosynthesis: Multifunctional enzyme: Phosphopantetheine;
 KM Hydroxylase; Oxidoreductase; Transferase; Lyase; NADP;
 KM Pyridoxal phosphate.
 FT DOMAIN 429 817 BETA-KETOACYL SYNTHASE.
 FT DOMAIN 1629 1857 ACYL AND MALONYL TRANSFERASES.
 FT DOMAIN 1858 2113 ENOYL REDUCTASE.
 FT BETA-KETOACYL REDUCTASE.

```

FT DOMAIN 2118 2174 ACYL CARRIER (ACP).
FT DOMAIN 2202 2505 THIOESTERASE.
FT ACT_SITE 161 161 BETA-KETOACYL SYNTHASE (BY SIMILARITY).
FT ACT_SITE 581 581 MALONYLTRANSFERASE (BY SIMILARITY).
FT NP_BIND 1658 1682 NADP (ER).
FT BINDING 1658 1688 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT NP_BIND 1765 1780 NADP (K4).
FT BINDING 2151 2151 PHOSPHOPANETHENINE (BY SIMILARITY).
FT ACT_SITE 2302 2302 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 2475 2475 THIOESTERASE (BY SIMILARITY).
FT ACT_SITE 878 878 BETA-HYDROXYACYL DEHYDRATASE (BY SIMILARITY).

Query Match 27.1%; Score 45; DB 1; Length 2505;
Best Local Similarity 26.3%; Pred. No. 2.8e+02;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SSFYGLFGQHPDYVYLMPEP 25
Db 40 WKAGLGLPKRSGKLKLSKFDASFQGVHPKQAHMTP 77

RESULT 33
AATM_BOVIN
ID AATM_BOVIN STANDARD; PRT; 430 AA.
AC P12344;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
GN GOT2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RA NCB1
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Palmisano A.;
RL Submitted (OCT-1993) to the EMBL/GenBank/DBJ databases.

[2]
RP SEQUENCE OF 30-41.
RX MEDLINE=79191877; PubMed=446759;
RA Capasso S., Garzillo A.M., Marino G., Mazarella L., Pucci P.,
RA Sanna G.;
RT "Mitochondrial bovine aspartate aminotransferase. Preliminary
sequence and crystallographic data."
RL FEBS Lett. 101:351-354(1979).

-1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
L-glutamate.
-1- COFACTOR: PYRIDOXAL PHOSPHATE.
-1- SUBUNIT: HOMODIMER.
-1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
ONE AND A MITOCHONDRIAL ONE.
-1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: 225466; CA80960.1;
DR PIR: A14290; A14290.
DR HSSP: P00508; 7AAT.
DR InterPro: IPR004839; Aminotransf./2.
DR InterPro: IPR000796; Aspttransf_sub.
DR InterPro: IPR004838; Nhltransf_1.
DR Pfam: PF00155; aminotran_1_2; 1.
DR PRINTS: PR00799; TRANSAMINASE.
DR PROSITE: PS00105; AA TRANSFER CLASS 1; 1.
KW Transfere; Aminotransferase; Pyridoxal phosphate; Mitochondrion;
KW Transit peptide.
FT TRANSIT 1 29 MITOCHONDRION.
FT CHAIN 30 430 ASPARTATE AMINOTRANSFERASE.
FT BINDING 279 279 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 430 AA; 47513 MW; 16DDF4753820355AA CRC64;

Query Match 26.8%; Score 44.5; DB 1; Length 430;
Best Local Similarity 44.0%; Pred. No. 52;
Matches 11; Conservative 4; Mismatches 9; Indels 1; Gaps 1;

QY 3 RSGSFVGLFGQHPDYVYLMPEP 27
Db 139 RIGASFGLRFFKFRDVF-LPKPTM 162

RESULT 34
AATM_HUMAN
ID AATM_HUMAN STANDARD; PRT; 430 AA.
AC P00505; Q9BNA3;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Aspartate aminotransferase, mitochondrial precursor (EC 2.6.1.1)
DE (Transaminase A) (Glutamate oxaloacetate transaminase-2).
GN GOT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RA NCB1
RP SEQUENCE FROM N.A.
RX MEDLINE=89087454; PubMed=3207426;
RA Pol S., Bousquet-Lemerclier B., Pave-Preux M., Pawlak A., Nalpas B.,
RA Berthelot P., Hanoune J., Barouki R.;
RT "Nucleotide sequence and tissue distribution of the human
mitochondrial aspartate aminotransferase mRNA."
RL Biochem. Biophys. Res. Commun. 157:1305-1315(1988).

[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Muscle;
RA Strausberg R.;
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.

[3]
RP SEQUENCE OF 30-430.
RX MEDLINE=86026367; PubMed=4052435;
RA Martini F., Angelaccio S., Barra D., Pascarella S., Maras B.,
RA Doonan S., Bossa F.;
RT "The primary structure of mitochondrial aspartate aminotransferase
from human heart."
RL Biochim. Biophys. Acta 832:46-51(1985).

-1- CATALYTIC ACTIVITY: L-aspartate + 2-oxoglutarate = oxaloacetate +
L-glutamate.
-1- COFACTOR: PYRIDOXAL PHOSPHATE.
-1- SUBUNIT: HOMODIMER.
-1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO ISOZYMES: A CYTOPLASMIC
ONE AND A MITOCHONDRIAL ONE.
-1- SIMILARITY: BELONGS TO CLASS-I OF PYRIDOXAL-PHOSPHATE-DEPENDENT
AMINOTRANSFERASES.

-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```


Query Match 26.5%; Score 44; DB 1; Length 170;
Best Local Similarity 37.5%; Pred. No. 23;
Matches 12; Conservative 2; Mismatches 10; Indels 8; Gaps 1;

QY 4 SSSFFVGLPFGQHPDVFYLM 27
DB 68 TSSSLKGFIDKGVDAKYNINHLQGLFTDPBW 99

RESULT 37
RCSA_SALTY
ID RCSA_SALTY STANDARD; PRT: 207 AA.
AC 056083;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Colanic acid capsular biosynthesis activation protein A.
GN RCSA OR STY2190.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=601;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Ty2;
RX MEDLINE=96198173; PubMed=8626298;
RA Vriologues I., Maxin H., Ecobichon C., Lee J.O., Popoff M.Y.;
RT "Characterization of the rcsA and rcsB genes from Salmonella typhi:
rCSB through tvaA is involved in regulation of Vi antigen
synthesis";
RT J. Bacteriol. 178:1691-1698(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-CT18;
RX MEDLINE=21534947; PubMed=11677608;
RA Parrilli J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrett B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhimurium CT18";
RL Nature 413:848-852(2001).
CC -1- FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE
SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION
OF THE GENES FOR CAPSULE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE LUXR/OHPA FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X87687; CA61019.1; -
DR EMBL: AL627272; CAD05730.1; -
DR InterPro: IPR000792; HTH_LuxR.
DR Pfam: PF00196; Gede; 1.
DR PRINTS: PR00038; HTHLuxR.
DR PRODOM: PD000307; HTH_LuxR; 1.
DR SMART: SM00421; HTH_LuxR; 1.
DR PROSITE: PS00622; HTH_LuxR_FAMILY; 1.
KW Sensory transduction; Transcription regulation; DNA-binding;
KW Activator; Complete proteome.
FT DNA_BIND 155 174 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23274 MW; 4DC34F9501C66999 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 207;
Best Local Similarity 35.0%; Pred. No. 29;
Matches 7; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY 4 SSSFFVGLPFGQHPDVFYLM 23
DB 63 SDSQIKQIINOHDPDTFTV 82

RESULT 38
RCSA_SALTY
ID RCSA_SALTY STANDARD; PRT: 207 AA.
AC P54698;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Colanic acid capsular biosynthesis activation protein A.
GN RCSA OR SPM1982.
OS Salmonella typhimurium.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-LT2 / SCSG1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., All J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Gwral N., Mulvaney E.,
RA Ryan E., Sun H., Flores L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2";
RT Nature 413:852-856(2001).
RN [2]
RP SEQUENCE OF 1-65 FROM N.A.
RC STRAIN-LT2;
RX MEDLINE=9746436; PubMed=9324257;
RA Chisholm K., Fan F., Schoenheits G.J., Kihara M., Macnab R.M.;
RT "The FljO, FljP, FljQ, and FljR proteins of Salmonella typhimurium:
putative components for flagellar assembly";
RT J. Bacteriol. 179:6092-6099(1997).
CC -1- FUNCTION: POSITIVE REGULATOR OF CAPSULAR POLYSACCHARIDE
SYNTHESIS. RCSA AND RCSB FORM A COMPLEX TO PROMOTE TRANSCRIPTION
OF THE GENES FOR CAPSULE SYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE LUXR/OHPA FAMILY OF TRANSCRIPTIONAL
REGULATORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE008788; AAU20894.1; -
DR EMBL: LA9021; AAB81322.1; -
DR STGene: SG10580; rcsA.
DR InterPro: IPR000792; HTH_LuxR.
DR Pfam: PF00196; Gede; 1.
DR PRODOM: PD000307; HTH_LuxR; 1.
DR PROSITE: PS00622; HTH_LuxR_FAMILY; 1.
KW Sensory transduction; Transcription regulation; DNA-binding;
KW Activator; Complete proteome.
FT DNA_BIND 155 174 H-T-H MOTIF (BY SIMILARITY).
SQ SEQUENCE 207 AA; 23302 MW; 4DC34F8A11C66999 CRC64;

Db 63 SDSQIKIINQHPDLEIFV 82

```

RESULT 39
ID Y1E_ECOLI STANDARD; PRT; 301 AA.
AC P32667;
DT 01-OCT-1993 (Rel. 27, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical transport protein y1E.
GN Y1E OR B3943.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=94089392; PubMed=8265357;
RA Blattner F.R., Burland V.D., Plunkett G. III, Sofia H.J.,
RA Daniels D.L.;
RT "Analysis of the Escherichia coli genome. IV. DNA sequence of the
RT region from 89.2 to 92.8 minutes."
RL Nucleic Acids Res. 21:5408-5417(1993).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: BELONGS TO THE EMA TRANSPORTER FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U00006; AAC43049.1; ALT_INT.
DR EMBL; AE000468; AAC76925.1; ALT_INT.
DR EcoGene; EG11902; Y1E.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 9 29 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 72 92 POTENTIAL.
FT TRANSMEM 99 119 POTENTIAL.
FT TRANSMEM 128 148 POTENTIAL.
FT TRANSMEM 151 171 POTENTIAL.
FT TRANSMEM 186 206 POTENTIAL.
FT TRANSMEM 219 239 POTENTIAL.
FT TRANSMEM 246 266 POTENTIAL.
FT TRANSMEM 269 289 POTENTIAL.
SQ SEQUENCE 301 AA; 32865 MW; 1A822058FDA28696 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 301;
Best Local Similarity 27.6%; Pred. No. 43;
Matches 8; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 SWRSSSFVQGLFGQHPDVFYLMPEAMHV 29
Db 163 SWGASATVAKRLYARHPVDLSLTJWQM 191

RESULT 40
ID YE09_SYNT3 STANDARD; PRT; 326 AA.
AC P73594;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical WD-repeat protein sir1409.
GN SIR1409.

```

```

OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions."
RL DNA Res. 3:109-136(1996).
CC -1- SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D90907; BAA17638.1; -
DR InterPro; IPR001680; WD40.
DR Pfam; PF00400; WD40; 7.
DR PRINTS; PR00320; GPROTEINRPT.
DR PRODOM; PD000018; WD40; 1.
DR SMART; SM00320; WD40; 7.
DR SMART; PS00678; WD_REPEATS_1; 2.
DR PROSITE; PS00682; WD_REPEATS_2; 4.
DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
KW Hypothetical protein; Repeat; WD repeat; Complete proteome.
FT REPEAT 47 77 WD 1.
FT REPEAT 88 118 WD 2.
FT REPEAT 129 159 WD 3.
FT REPEAT 169 199 WD 4.
FT REPEAT 210 240 WD 5.
FT REPEAT 252 282 WD 6.
SQ SEQUENCE 326 AA; 35759 MW; 5D83A2403DE1163D CRC64;

Query Match 26.5%; Score 44; DB 1; Length 326;
Best Local Similarity 42.1%; Pred. No. 46;
Matches 8; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 2 WRSSSFVQGLFGQHPDVF 20
Db 76 WTKGEMTLGQLGQKPMF 94

RESULT 41
ID ALR_DEIRA STANDARD; PRT; 351 AA.
AC Q9VE3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alanine racemase (EC 5.1.1.1).
GN ALR OR DR1086.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RL;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,

```

RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 CC "Genome sequence of the radioresistant bacterium *Deinococcus*
 RT *radiodurans* R1.";
 RL Science 286:1571-1577(1999).
 CC -1- FUNCTION: Provides the D-alanine required for cell wall
 CC biosynthesis (By similarity).
 CC -1- CATALYTIC ACTIVITY: L-alanine = D-alanine.
 CC -1- COFACTOR: Pyridoxal phosphate (By similarity).
 CC -1- PATHWAY: Along with D-alanine-D-alanine ligase, it makes up the
 CC D-alanine branch of the peptidoglycan biosynthetic route.
 CC -1- SIMILARITY: BELONGS TO THE ALANINE RACEMASE FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE001958; AAF10657.1; ALT-INIT.
 CC HSSP: P10724; 1BD0.
 DR TIGR: DR1086;
 DR InterPro: IPR000821; Ala-racemase.
 DR Pfam: PF00842; Ala-racemase. 1.
 DR PRINTS: PR00992; ALARACEMASE.
 DR TIGRPM: TIGR00492; alt. 1.
 DR PROSITE: PS00395; ALANINE_RACEMASE. 1.
 KW Isomerase; Pyridoxal phosphate; Cell wall; Peptidoglycan synthesis;
 KW Complete proteome.
 FT ACT_SITE 34 CATALYTIC BASE SPECIFIC TO D-ALANINE (BY
 FT SIMILARITY).
 FT ACT_SITE 248 CATALYTIC BASE SPECIFIC TO L-ALANINE (BY
 FT SIMILARITY).
 FT BINDING 34 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 FT SEQUENCE 351 AA; 37192 MW; F924ABD662E2FPA05 CRC64;
 SQ
 Query Match 26.5%; Score 44; DB 1; Length 351;
 Best Local Similarity 50.0%; Pred. No. 50;
 Matches 9; Conservative 2; Mismatches 7; Indels 0; Gaps 0;
 QY 5 GGSFVQGLFGQHPDYFYL 22
 DB 41 GMELVGRLLAAHPDVWGL 58
 RESULT 42
 SEBL_XENLA
 ID SEBL_XENLA STANDARD; PRT; 481 AA.
 AC P58003;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 16-OCT-2001 (Rel. 40; Last annotation update)
 DE Sestrin 1 (p53-regulated protein PA26) (XPA26).
 GN SEST1 OR PA26.
 OS *Xenopus laevis* (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodinae; *Xenopus*.
 OX NCBI_TaxID=8355;
 RN
 RP [1]
 RP SEQUENCE FROM N.A.
 RP MEDLINE=21095182; PubMed=1165487;
 RA Hikasa H., Taira M.;
 RA "A *Xenopus* homolog of a human p53-activated gene, PA26, is
 RT specifically expressed in the notochord.";
 RL Mech. Dev. 100:309-312(2001).
 CC -1- SUBCELLULAR LOCATION: Nuclear.
 CC -1- DEVELOPMENTAL STAGE: AT ZYGOTIC STAGE; SPECIFICALLY EXPRESSED IN
 CC THE NOTOCHORD. MATERNAL TRANSCRIPTS ARE DETECTED AT CLEAVAGE
 CC STAGES AND REDUCED DURING GASTRULATION.

CC -1- SIMILARITY: BELONGS TO THE SESTRIN FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AB048259; BAB33008.1;
 KW Nuclear protein.
 KW SEQUENCE 481 AA; 55459 MW; 8C175801BBFAFE2D CRC64;
 SQ
 Query Match 26.5%; Score 44; DB 1; Length 481;
 Best Local Similarity 36.7%; Pred. No. 70;
 Matches 11; Conservative 6; Mismatches 9; Indels 4; Gaps 2;
 QY 1 SMRS-GGSFVQGLFGQHPDYFVMEPAWV 29
 DB 344 SWEHGYSLVNRLY---FDVGQLDDEKFTI 370
 RESULT 43
 UVRC_HAEIN
 ID UVRC_HAEIN STANDARD; PRT; 609 AA.
 AC P4485;
 DT 01-NOV-1995 (Rel. 32; Created)
 DT 01-NOV-1995 (Rel. 32; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Excinuclease ABC subunit C.
 GN UVRC OR HI0057.
 OS *Haemophilus influenzae*.
 CC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 CC *Haemophilus*.
 OX NCBI_TaxID=727;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=Rd / KM20 / ATCC 51907;
 RP MEDLINE=95350630; PubMed=7542800;
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
 RA Kevlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
 RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geochagen N.S.M.,
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
 RA Venter J.C.;
 RA "Whole-genome random sequencing and assembly of *Haemophilus influenzae*
 RT Rd.";
 RL Science 269:496-512(1995).
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT
 CC CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS
 CC PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). ATTACHES TO THE
 CC UVRA-UVRB COMPLEX, DISPLACING UVRA, AND THE DAMAGED DNA STRAND IS
 CC NICKED ON BOTH SIDES OF THE DAMAGED SITE (BY SIMILARITY).
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNITS: UVRA, UVRB AND UVRC.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -1- SIMILARITY: CONTAINS 1 UVR DOMAIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U32691; AAC21735.1;
 DR HSSP: P07025; IE52.
 DR TIGR: HI0057;
 DR InterPro: IPR003583; HHH_1.

DR InterPro: IPR000445; Hh.
 DR InterPro: IPR001943; UVR/C.
 DR InterPro: IPR004791; UVR/C.
 DR InterPro: IPR001162; UVR/C.
 DR InterPro: IPR000305; UVR/N.
 DR Pfam: PF00633; HhH. 2.
 DR Pfam: PF01541; Excl_endo_N; 1.
 DR Pfam: PF02151; UVR; 1.
 DR ProDom: PD005870; UVR_C; 1.
 DR SMART: SM00465; GYRC; 1.
 DR SMART: SM00278; HhH1; 1.
 DR TIGRFAMs: TIGR00194; uvrC; 1.
 DR PROSITE: PS0151; UVR; 1.
 KM SOS response: Excision nuclease: DNA repair; Complete proteome.
 FT DOMAIN 201 236 UVR.
 SQ SEQUENCE 609 AA; 69609 MW; 049DD9212349017A CRC64;

Query Match 26.5%; Score 44; DB 1; Length 609;
 Best Local Similarity 41.2%; Pred. No. 89;
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

DB 114 SYRGSSEFGVQLFGQHP 17
 114 SYRGSSEFGVQLFGQHP 130

RESULT 44
 S6AB_MOUSE STANDARD; PRT; 627 AA.
 ID S6AB_MOUSE
 AC P31650;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-OCT-2001 (Rel. 40, Last annotation update)
 DE Sodium- and chloride-dependent GABA transporter 4 (GAT4).
 GN SLC6A11 OR GABT4 OR GAT-4 OR GAT4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93111969; PubMed=8420981;
 RA Liu Q.-R., Lopez-Coquera B., Mandiyan S., Nelson H., Nelson N.;
 RT "Molecular characterization of four pharmacologically distinct gamma-
 RT aminobutyric acid transporters in mouse brain.";
 RL J. Biol. Chem. 268:2106-2112(1993).
 CC -!- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS. HAS A KM
 CC OF 0.8 UUM FOR GABA; CAN ALSO TRANSPORT BETA-ALANINE (KM=99 UUM)
 CC AND TAURINE (KM=1.4 UUM).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: BRAIN.
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
 CC FAMILY (SNP).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: L04662; -!- NOT_ANNOTATED_CDS.
 DR PIR: B44409; B44409.
 DR MGD: MGI:95630; Gabt4.
 DR InterPro: IPR000175; Na/ntran_symport.
 DR Pfam: PF00209; SNF; 1
 DR PRINTS: PR00176; NANEUSMPORT.
 DR ProDom: PD000448; Na/ntran_symport; 1.
 DR PROSITE: PS00610; NA_NEUROTRAN_SYM_1; 1.
 DR PROSITE: PS00754; NA_NEUROTRAN_SYM_2; 1.

DR PROSITE: PS50267; NA_NEUROTRAN_SYM_3; 1.
 KM Neurotransmitter transport; Transmembrane; Glycoprotein;
 KW Symport; Multigene family.
 FT DOMAIN 1 53
 FT TRANSMEM 54 74
 FT TRANSMEM 82 101
 FT TRANSMEM 126 146
 FT DOMAIN 147 220
 FT TRANSMEM 221 239
 FT TRANSMEM 248 265
 FT TRANSMEM 301 318
 FT TRANSMEM 330 351
 FT TRANSMEM 384 403
 FT TRANSMEM 433 451
 FT TRANSMEM 468 488
 FT TRANSMEM 509 528
 FT TRANSMEM 548 566
 FT DOMAIN 567 627
 FT CARBOHYD 182 182
 FT CARBOHYD 185 185
 FT CARBOHYD 193 193
 SQ SEQUENCE 627 AA; 69888 MW; E6D4E45FC92ACB7 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 627;
 Best Local Similarity 26.5%; Pred. No. 92;
 Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

DB 117 WRVRCPLFEGIGYATQVIEAHNLVYIIILAMAI 150

RESULT 45
 S6AB_RAT STANDARD; PRT; 627 AA.
 ID S6AB_RAT
 AC P31647;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUL-1998 (Rel. 36, Last annotation update)
 DE Sodium- and chloride-dependent GABA transporter 3.
 GN SLC6A11 OR GABT3 OR GAT-3 OR GAT-B.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Brain;
 RX MEDLINE=93016029; PubMed=1400419;
 RA Borden L.A., Smith K.E., Hartig P.R., Brancheck T.A., Weinschenk R.L.;
 RT "Molecular heterogeneity of the gamma-aminobutyric acid (GABA)
 RT transport system. Cloning of two novel high affinity GABA
 RT transporters from rat brain.";
 RL J. Biol. Chem. 267:21098-21104(1992).
 CC -!- FUNCTION: TERMINATES THE ACTION OF GABA BY ITS HIGH AFFINITY
 CC SODIUM-DEPENDENT REUPTAKE INTO PRESYNAPTIC TERMINALS.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: BRAIN AND RETINA.
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:NEUROTRANSMITTER SYMPORTER
 CC FAMILY (SNP).
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial


```

RT gastricpathogen Helicobacter pylori." ;
RL Nature 397:176-180(1999).
CC CC
CC -I- CATABOLIC ACTIVITY: ATP + pyruvate + H(2)O = AMP +
CC phosphoenolpyruvate + phosphate.
CC -I- PATHWAY: ESSENTIAL STEP IN GLUCONEOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -I- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb.ch/announce/
CC or send an email to license@isb.slb.ch).
CC CC
DR EMBL; AE001450; AAD05690.1.
DR InterPro; IPR000121; pep_utilizers.
DR InterPro; IPR000121; PPK_N_term.
DR Pfam; PF003391; pep_utilizers; 1.
DR Pfam; PF01326; PPK_Nterm; 1.
DR Pfam; PF02896; pep_utilizers.C; 1.
DR ProDom; PD000940; pep_utilizers; 1.
DR PROSITE; PS00370; pep_ENZYMES_PHOS_SITE; 1.
DR PROSITE; PS00742; pep_UTILIZES_2; 1.
DR TRANSFERASE; kinase; ATP-binding; Phosphorylation, complete proteome.
KW MOD_RES 430 430 PHOSPHORYLATION (BY SIMILARITY).
FT SEQUENCE 812 AA; 91291 MW; A088D459B3E47512 CRC64;
SQ
Query Match 26.5%; Score 44; DB 1; Length 812;
Best local Similarity 36.0%; Fred. NO. 1.2e+02;
Matches 9; Conservative 5; Mismatches 9; Indels 2; Gaps 1;
OY 1 SWRSSGSEFGQFQGHPVFLYLMPE 25
DB 220 AMGLGNENVGGIT--NPDEFFYFKP 242
| | : | : || | : | : |
RESULT 48
PPSA_HELPEY STANDARD: PRt; 812 AA.
AC P56070;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Phosphoenolpyruvate synthase (EC 2.7.9.2) (Pyruvate/water dikinase)
DE (PEP synthase).
DE PPSA OR HP0121.
GN GN HP0121.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ANCG 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Nelson K.M., Quackenbush J., Zhou L., Kirkness E.F., Dougherty B.A.,
RA Loftus B., Richardson D., Dodson R., Khairak H.G., Glodek A.,
RA McEnenny K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utechtack T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karpi P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori." ;
RL Nature 388:539-547(1997).
CC -I- CATABOLIC ACTIVITY: ATP + pyruvate + H(2)O = AMP +
CC phosphoenolpyruvate + phosphate
CC -I- PATHWAY: ESSENTIAL STEP IN GLUCONEOGENESIS WHEN PYRUVATE AND
CC LACTATE ARE USED AS A CARBON SOURCE.
CC -I- SIMILARITY: BELONGS TO THE PEP-UTILIZING ENZYMES FAMILY.
```

```
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; AE000534; AAD07191.1; -
DR      TIGR; HP0121; -
DR      InterPro; IPR000121; PEP_utilizers.
DR      InterPro; IPR002192; PPDK_N term.
DR      Pfam; PF00391; PEP-utilizers; 1.
DR      Pfam; PF01326; PEP-N term; 1.
DR      Pfam; PF02896; PEP-utilizers; 1.
DR      Pfam; PD000940; PEP-utilizers; 1.
DR      PROSITE; PS00370; PEP_ENZYMES_PHOS_SITE; 1.
DR      PROSITE; PS00742; PEP_ENZYMES_2; 1.
DR      Transferrase; Kinase; ATP-binding; Phosphorylation; Complete proteome.
KT      MOD_RES     430      PHOSPHORYLATION (BY SIMILARITY).
FT      SEQUENCE   812 AA;  91273 MW;  560660703B134AA3 CRC64;
SQ
Query Match                                26.5%; Score 44; DB 1; Length 812;
Best Local Similarity                      36.0%; Pred. No. 1.2e+02;
Matches    9; Conservative    5; Mismatches    9; Indels    2; Gaps    1;
QY      1 SRRSSSFVGQLFGQHDPVFLMEP 25
          :|::||::||::||::||
DB      220 AWGLGENVGGIT--NPDEIFYFKP 242

RESULT 49
VG26_BPMD2 ID VG26_BPMD2 STANDARD: PRT: 836 AA.
AC      O64220; STANFORD;
DT      15-DEC-1998 (Rel. 37, Created)
DI      15-DEC-1998 (Rel. 37, Last sequence update)
DE      15-DEC-1998 (Rel. 37, Last annotation update)
DR      Minor tail protein GP26.
DS      26.
GS      Mycobacteriophage D29.
OC      Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae;
OC      Unclassified Siphoviridae.
OX      NCBI_TaxID=28369;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=98300335; Pubmed=9636706;
RA      Ford M.E., Sarkis G.J., Belanger A.E., Hendrix R.W., Hatfull G.F.;
RT      "Genome structure of mycobacteriophage D29: Implications for phage
RL      evolution.";
RL      J. Mol. Biol. 279:143-164(1998).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      The European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL; AF022214; AAC18467.1; -
DR      INTN_MET 0 BY SIMILARITY.
SQ      SEQUENCE   836 AA;  86572 MW;  528D12ED0FCB92D6 CRC64;
QY      Query Match                                26.5%; Score 44; DB 1; Length 836;
          Best Local Similarity                      57.9%; Pred. No. 1.2e+02;
          Matches    11; Conservative    1; Mismatches    3; Indels    4; Gaps    1;
          QY      1 SRRSG----SFRVGOLEQG 15
                |||||
DB      560 SWSGCGOQVDFVGOGLPKK 578
```

```

RESULT 50
POLR ASGVP
ID POLR ASGVP STANDARD: PRT: 2105 AA.
AC P36309;
DT 01-JUN-1994 (rel. 29, Created)
DT 01-JUN-1994 (rel. 29, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Genome polyprotein [contains: RNA replicase (EC 2.7.7.48); Helicase;
DE Coat protein].
OS Apple stem grooving virus (strain P-209) (ASGV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Capillovirus.
OX NCBI_TaxID=36402;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93033164; Pubmed=1413530;
RA Yoshikawa N., Sasaki E., Kato M., Takahashi T.;
RT "The nucleotide sequence of apple stem grooving capillovirus genome.";
RL Virology 191:98-105(1992).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -1- PTM: THE N-TERMINAL OF THE COAT PROTEIN IS BLOCKED.
CC -1- MISCELLANEOUS: THE COAT PROTEIN IS LOCATED IN THE CARBOXY-TERMINAL
CC REGION OF THIS POLYPROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14995; BA03639.1; -.
DR PIR: A44059; A44059.
DR MEROPS: C35.001; -.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Rfam: RF01443; Viral_helicase1; 1.
KW RNA-directed RNA polymerase; Transferase; Polyprotein; ATP-binding;
KW Coat protein; Helicase.
FT NP_BIND 781 788 ATP (POTENTIAL).
FT DOMAIN 1364 1453 POLYMERASE ACTIVE SITE.
SQ SEQUENCE 2105 AA; 241240 MW; 331291A5FA137131 CRC64;

Query Match 26.5%; Score 44; DB 1; Length 2105;
Best Local Similarity 33.3%; Pred. No. 3.3e+02;
Matches 9; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

```

Search completed: February 20, 2003, 13:32:24
 Job time : 41.4286 secs


```

966 40 24.1 513 8 09T199 09T199 orcuttia ca
967 40 24.1 513 8 09T197 09T197 sporobolus
968 40 24.1 513 8 09T196 09T196 monodia sti
969 40 24.1 513 8 09T195 09T195 plectrachne
970 40 24.1 513 8 09T194 09T194 papophorum
971 40 24.1 513 8 09T193 09T193 eragrostis
972 40 24.1 513 8 09T192 09T192 fingerhuthi
973 40 24.1 513 8 09T191 09T191 uniola pani
974 40 24.1 513 8 09T190 09T190 arundo dona
975 40 24.1 513 8 09T189 09T189 denthonia s
976 40 24.1 513 8 09T188 09T188 centropodia
977 40 24.1 513 8 09T187 09T187 molinia cae
978 40 24.1 513 8 09T186 09T186 chasmanthiu
979 40 24.1 513 8 09T185 09T185 lophatherum
980 40 24.1 513 8 09T184 09T184 orthocladia
981 40 24.1 513 8 09T183 09T183 tristrachya
982 40 24.1 513 8 09T182 09T182 digitaria s
983 40 24.1 513 8 09T181 09T181 panicum cap
984 40 24.1 513 8 09T180 09T180 zoyzia macr
985 40 24.1 513 8 09T179 09T179 eragrostis
986 40 24.1 513 8 09T178 09T178 kengia song
987 40 24.1 513 8 09T177 09T177 sporobolus
988 40 24.1 513 8 09T176 09T176 eragrostis
989 40 24.1 513 8 09T175 09T175 eragrostis
990 40 24.1 513 8 09T174 09T174 mulienbergi
991 40 24.1 513 8 09T173 09T173 tridens bra
992 40 24.1 513 8 09T172 09T172 aegopogon c
993 40 24.1 513 8 09T171 09T171 melanocench
994 40 24.1 513 8 09T170 09T170 brachyachne
995 40 24.1 513 8 09T169 09T169 chloris can
996 40 24.1 513 8 09T168 09T168 chloris pyc
997 40 24.1 513 8 09T167 09T167 chloris tra
998 40 24.1 513 8 09T166 09T166 cynodon tra
999 40 24.1 513 8 09T165 09T165 enteropogon
1000 40 24.1 513 8 09T164 09T164 oxychloris

```

ALIGNMENTS

```

RESULT 1
ID 09Y5R3 PRELIMINARY; PRT: 386 AA.
AC 09Y5R3:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand
DE sulfotransferase GST-3).
GN GST3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RX MEDLINE=99764336; PubMed=10330415;
RA Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R.,
RA Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.;
RA "Sulfotransferases of two specificities function in the reconstitution
RT of high endothelial cell ligands for L-selectin.";
RT J. Cell Biol. 145:899-910(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
RA Rosen S.D.;
RA "Chromosomal Localization and Genomic Organization for the
RT Galactose/N-Acetylglucosamine/N-Acetylglucosamine 6-O-
RT Sulfotransferase Gene Family.";
RT Glycobiology 0:0-0(2001).
RN [3]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=11439191; PubMed=11439191;
RA Yeh J.-C., Hiraoka N., Petryniak B., Nakayama J., Ellies L.G.,
RA Raduka D., Hindsgraff O., March J.D., Lowe J.B., Fakuda M.;
RT "Novel sulfated lymphocyte homing receptors and their control by a
RT core1 extension beta1,3-N-acetylglucosaminyltransferase.";
RT Cell 105:957-969(2001).
RL Cell 105:957-969(2001).
DR EMBL: AF131235; AAD33015.1; -.
DR EMBL: AF280088; AAG48246.1; -.
DR EMBL: AF149783; AAK48417.1; -.
DR INTERP: IPR001092; ELH_Dasic.
DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
DR Lectin: Selectin; Transferase.
SQ SEQUENCE 386 AA; 4513 MW; 0C3B4022417143A CRC64;

```

```

Query Match 100.0%; Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2,7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SMRSGSFVQGLFGQHPDVFYLMPPAMHV 29
DB 50 SMRSGSFVQGLFGQHPDVFYLMPPAMHV 78

```

```

RESULT 2
ID 09R1I1 PRELIMINARY; PRT: 388 AA.
AC 09R1I1:
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE L-selectin ligand sulfotransferase.
GN CHST4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9361934; PubMed=10435581;
RA Hiraoka N., Petryniak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,
RA Izawa D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.;
RT "A novel, high endothelial venule-specific sulfotransferase expresses
RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.";
RL Immunity 11:79-89(1999).
DR EMBL: AF109155; AAD45579.1; -.
DR MGI: MGI:1349479; Chst4.
KW Lectin; Selectin; Transferase.
SQ SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884AEE CRC64;

```

```

Query Match 100.0%; Score 166; DB 11; Length 388;
Best Local Similarity 100.0%; Pred. No. 2,7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 SMRSGSFVQGLFGQHPDVFYLMPPAMHV 29
DB 49 SMRSGSFVQGLFGQHPDVFYLMPPAMHV 77

```

```

RESULT 3
ID 09W0E5 PRELIMINARY; PRT: 388 AA.
AC 09W0E5:
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase (Chondroitin
DE 6/keratan) sulfotransferase 4).
GN CHST4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```



```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=TONSIL;
RX MEDLINE=99264336; PubMed=10330415;
RA Bistrup A., Bhakta S., Lee J.-K., Belov Y.Y., Gunn M.D., Zuo F.-R.,
RA Huang C.-C., Kannagi R., Rosen S.D., Hemmerich S.,
RT "Sulfoltransferases of two specificities function in the reconstitution
RT of high endothelial cell ligands for L-selectin.";
RL J. Cell Biol. 145:899-910(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=TONGUE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itch M., Ishi Y.,
RA Akakawa T., Hara A., Fukunishi Y., Komro H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitlaker C., Wilming L.,
RA Wyshak-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AF131326; AAD33016.1; -
DR EMBL: AK009113; BAB26078.1; -
DR MGD: MGI:1349479; Chst4.
RW Transferase.
SQ SEQUENCE 388 AA; 44694 MW; D1E9D7796DF857AD CRC64;

Query Match 100.0%; Score 166; DB 11; Length 388;
Best Local Similarity 100.0%; Pred. No. 2, 7e-15;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 29
DB 49 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 77

RESULT 4
Q9GZS9 PRELIMINARY; PRT; 395 AA.
AC Q9GZS9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfoltransferase GST-4beta (Corneal N-
DE acetylglucosamine-6-O-sulfoltransferase).
GN GST4BETA OR CHST6.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RX [1]
RP SEQUENCE FROM N.A.
RA Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
RA Rosen S.D.;
RT "Chromosomal localization and genomic organization for the
RT galactose/N-acetylgalactosamine/N-acetylglucosamine 6-O-
RT sulfoltransferase gene family.";
RL Glycobiology 0:0-0(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20472330; PubMed=11017086;

```

```

RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Fujiwara T.,
RA Nakamura T., Dota A., Kawasaki S., Inoue Y., Meeda N., Yamamoto S.,
RA Ozaki K., Kinoshita S., Shimomura Y., Tanigami A., Fukuda M.N.;
RT "Macular corneal dystrophy type I and type II are caused by distinct
RT mutations in a new sulfoltransferase gene.";
RL Nat. Genet. 26:237-241(2000).
DR EMBL: AF280086; AAG48244.1; -
DR EMBL: AF219990; AAG26325.1; -
DR EMBL: AF219991; AAG26327.1; -
RW Transferase.
SQ SEQUENCE 395 AA; 44098 MW; 433CA60248A48P67 CRC64;

Query Match 96.4%; Score 160; DB 4; Length 395;
Best Local Similarity 96.6%; Pred. No. 2e-14;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 29
DB 48 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 76

RESULT 5
Q9GZS9 PRELIMINARY; PRT; 395 AA.
AC Q9GZS9;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfoltransferase.
GN CHST5 OR I-GLCNAC-6-ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RX [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=INTESTINE;
RX MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
RT "Cloning and characterization of a mammalian N-acetylglucosamine-6-
RT sulfoltransferase that is highly restricted to intestinal tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL: AF176841; AAD56003.1; -
DR EMBL: AF176840; AAD56002.1; -
DR MGD: MGI:1931825; Chst5.
RW Transferase.
SQ SEQUENCE 395 AA; 44537 MW; 3FDF71E43ED383BE CRC64;

Query Match 96.4%; Score 160; DB 11; Length 395;
Best Local Similarity 96.6%; Pred. No. 2e-14;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 29
DB 48 SWRSGSFFVGLFGQHPDVFYLMPEAMHV 76

RESULT 6
Q9GZS9 PRELIMINARY; PRT; 411 AA.
AC Q9GZS9; Q9GZS9;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Intestinal GLCNAC-6-sulfoltransferase (Intestinal N-
DE acetylglucosamine-6-O-sulfoltransferase) (N-acetylglucosamine
DE 6-O-sulfoltransferase).
GN CHST5 OR I-GLCNAC-6-ST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OC NCBI_TaxID=9606;
RX [1]

```

RP SEQUENCE FROM N.A.
 RX MEDLINE=20472330; PubMed=11017086;
 RA Akama T.O., Nishida K., Nakayama J., Watanabe H., Ozaki K.,
 RA Nakamura T., Dots A., Kawasaki S., Inoue Y., Mada N., Yamamoto S.,
 RA Fujikawa T., Thonar E.J., Shimomura Y., Kinoshita S., Tanigami A.,
 RA Fukuda M.N.;
 RT "Macular corneal dystrophy type I and type II are caused by distinct
 RT mutations in a new sulphotransferase gene";
 RL Nat. Genet. 26:237-241(2000).
 RN [2]
 RP SEQUENCE OF 22-411 FROM N.A.
 RC TISSUE=INTESTINE;
 RX MEDLINE=942349; PubMed=1049128;
 RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerich S.;
 RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
 RT Sulfotransferase that is Highly Restricted to Intestinal Tissue";
 RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
 DR EMBL: AF246718; AAG28023.1; -;
 DR EMBL: AF218991; AAG28326.1; -;
 DR EMBL: AF176839; AAD56001.1; -;
 DR EMBL: AF176838; AAD56000.1; -;
 KM Transferase.
 SQ SEQUENCE 411 AA; 46160 MW; 97642D54BE26E06 CRC64;
 Query Match 94.6%; Score 157; DB 4; Length 411;
 Best Local Similarity 93.1%; Pred. No. 5.4e-14;
 Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 SMRSGSFGQLFGQHPDVFYLMPEAMHV 29
 Db 70 SMRSGSFGQLFGQHPDVFYLMPEAMHV 98
 RESULT 7
 Q9EP78 PRELIMINARY; PRT; 484 AA.
 AC Q9EP78;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE N-acetylglucosamine 6-O-sulfotransferase GST-5 (N-
 DE acetylglucosamine-6-O-sulfotransferase) (2600013M07R1k protein)
 DE (Carbohydrate (N-acetylglucosamine) sulfotransferase 7).
 GN CHST7 OR GST5 OR 2600013M07R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX PubMed=10956661;
 RA Bhakta S., Bartsch A., Bowman K.G., Kao W.M., Polsky I., Lee J.-K.,
 RA Cook B.N., Buehl R.E., Rosen S.D., Bertozzi C.R., Hemmerich S.;
 RT "Sulfation of N-acetylglucosamine by Chondroitin 6-Sulfotransferase 2
 RT (GST-5).";
 RL J. Biol. Chem. 275:40226-40234(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Uchimura K., Muramatsu T.;
 RT "Identification and molecular characterization of a cDNA encoding a
 RT novel N-acetylglucosamine-6-O-sulfotransferase";
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Toh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Oikido T., Furuno M., Aono H., Baldarelli R., Barch G.,
 RA Blake J., Boffelli D., Bojunga N., Carrinci P., de Bernaldo M.F.,
 RA Brownstein M.J., But C., Flechner C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Saeki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RX Strausberg R.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF280089; AAG48247.1; -;
 DR EMBL: AB040710; BAB13769.1; -;
 DR EMBL: AK011202; BAB27465.1; -;
 DR EMBL: BC019204; AAH19204.1; -;
 DR MGD: MGI:1891767; Chst7.
 KM Transferase.
 SQ SEQUENCE 484 AA; 54766 MW; 9B195537D7AB7193 CRC64;
 Query Match 80.7%; Score 134; DB 11; Length 484;
 Best Local Similarity 72.4%; Pred. No. 1.1e-10;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SMRSGSFGQLFGQHPDVFYLMPEAMHV 29
 Db 107 TWRGSSFLGELFNQHPDVFYLMPEAMHL 135
 RESULT 8
 Q9NS84 PRELIMINARY; PRT; 486 AA.
 AC Q9NS84;
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMblrel. 15, Last annotation update)
 DE Chondroitin 6-sulfotransferase-2.
 GN ChST-2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BRAIN;
 RX MEDLINE=20347214; PubMed=10781596;
 RA Kitagawa H., Fujita M., Ito N., Sugahara K.;
 RT "Molecular cloning and expression of a novel chondroitin 6-O-
 RT sulfotransferase";
 RL J. Biol. Chem. 275:21075-21080(2000).
 DR EMBL: AB037187; BAB03217.1; -;
 KM Transferase.
 SQ SEQUENCE 486 AA; 54294 MW; B524D1430B3C9075 CRC64;
 Query Match 80.7%; Score 134; DB 4; Length 486;
 Best Local Similarity 72.4%; Pred. No. 1.1e-10;
 Matches 21; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
 QY 1 SMRSGSFGQLFGQHPDVFYLMPEAMHV 29
 Db 109 TWRGSSFLGELFNQHPDVFYLMPEAMHL 137
 RESULT 9
 Q75667 PRELIMINARY; PRT; 486 AA.
 ID Q75667


```

RN [2]
RP SEQUENCE FROM N.A.
RA Peng T., Tabas I., Williams K.J.
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN:
RA MEDLINE=98070405; PubMed=9405439;
RX Fukuta M., Inazawa J., Torii T., Tsuzuki K., Shimada E., Habuchi O.;
RT "Molecular cloning and characterization of human keratan sulfate
   Gal-6-sulfotransferase."
RL J. Biol. Chem. 272:32321-32328(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=EMBRYONAL VEIN ENDOTHELIUM;
RX MEDLINE=99168906; PubMed=10049591;
RA Li X., Tedder T.F.;
RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
   endothelial cells: cDNA cloning, expression, and chromosomal
   localization."
RL Genomics 55:345-347(1999).
DR EMBL: U65637; AAC28776.1;
DR EMBL: AB003791; BAA24840.1;
DR EMBL: AF090137; AAD19878.1;
RW TRANSFERASE.
SQ SEQUENCE 411 AA; 46714 MW; F07D0A23B6338A09 CRC64;

Query Match
Best Local Similarity 69.9%; Score 116; DB 4; Length 411;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDYFLMEPAMHV 29
DB 70 RSGSFEVGLFGQHPDYFLMEPAMHV 29
|||||
ID 093403 PRELIMINARY; PRT; 441 AA.
AC 093403;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN NSIST.
OS Torpedo californica (Pacific electric ray).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC Elasmobranchii; Squalea; Hypnosqualea; Plistioraja; Batoidae;
OC Torpediniformes; Torpedinoidei; Torpedinidae; Torpedo.
OX NCBI_TaxID=7787;
RN [1]
RP SEQUENCE FROM N.A.
RA Nestuk M.A., Davis S., Yancopoulos G.D., Fallon J.R.;
RT "Expression cloning and characterization of NSIST, a novel
   sulfotransferase expressed by a subset of neurons and postsynaptic
   targets."
RL J. Neurosci. 0:0-0(1998).
DR EMBL: AF079675; AAC28491.1;
RW TRANSFERASE.
SQ SEQUENCE 441 AA; 50392 MW; 993B8B58743BA53 CRC64;

Query Match
Best Local Similarity 65.1%; Score 108; DB 13; Length 441;
Matches 18; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDYFLMEPAMHV 29
DB 107 RTGSSFEVGFENNDIFLFEPLMHV 133
|||||
ID 088199 PRELIMINARY; PRT; 472 AA.
AC 088199;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN CHST3 OR C6ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, AND 129; TISSUE=SPLEEN;
RX MEDLINE=98259835; PubMed=9597547;
RA Uchimura K., Kadomatsu K., Fan O.-W., Muramatsu H., Kurosawa N.,
   Kaname T., Yamamura K., Fukuta M., Habuchi O., Muramatsu T.;
RT "Mouse chondroitin 6-sulfotransferase: molecular cloning,
   characterization and chromosomal mapping."
RL Glycobiology 8:469-496(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638441; PubMed=11696535;
RA Uchimura K., Kadomatsu K., Nishimura H., Muramatsu H., Nakamura E.,
   Kurosawa N., Habuchi O., El-Fasakhay F.M., Yoshikai Y., Muramatsu T.;
RT "Functional analysis of the chondroitin 6-sulfotransferase gene in
   relation to lymphocyte subpopulations, brain development, and
   oversulfated chondroitin sulfates."
RL J. Biol. Chem. 277:1443-1450(2002).
DR EMBL: AB008937; BAA29054.1;
DR EMBL: AB062109; BAB72166.1;
DR MGD: MGI:185824; Chst3.
RW TRANSFERASE.
SQ SEQUENCE 472 AA; 53997 MW; F021147196D9D339 CRC64;

Query Match
Best Local Similarity 60.8%; Score 101; DB 11; Length 472;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDYFLMEPAMHV 29
DB 136 RTGSSFEVGFENNDIFLFEPLMHI 162
|||||
ID 090212 PRELIMINARY; PRT; 474 AA.
AC 090212;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Sciurognathi; Muridae; Murinae; Rattus.
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT "Chondroitin 6-sulfotransferase of rat sciatic nerve."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Liu J., Li X., Kwok C.-F., Tang S., Shum K.-Y.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF178689; AAD54386.2;

```

```

AC 088199;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN CHST3 OR C6ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J, AND 129; TISSUE=SPLEEN;
RX MEDLINE=98259835; PubMed=9597547;
RA Uchimura K., Kadomatsu K., Fan O.-W., Muramatsu H., Kurosawa N.,
   Kaname T., Yamamura K., Fukuta M., Habuchi O., Muramatsu T.;
RT "Mouse chondroitin 6-sulfotransferase: molecular cloning,
   characterization and chromosomal mapping."
RL Glycobiology 8:469-496(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21638441; PubMed=11696535;
RA Uchimura K., Kadomatsu K., Nishimura H., Muramatsu H., Nakamura E.,
   Kurosawa N., Habuchi O., El-Fasakhay F.M., Yoshikai Y., Muramatsu T.;
RT "Functional analysis of the chondroitin 6-sulfotransferase gene in
   relation to lymphocyte subpopulations, brain development, and
   oversulfated chondroitin sulfates."
RL J. Biol. Chem. 277:1443-1450(2002).
DR EMBL: AB008937; BAA29054.1;
DR EMBL: AB062109; BAB72166.1;
DR MGD: MGI:185824; Chst3.
RW TRANSFERASE.
SQ SEQUENCE 472 AA; 53997 MW; F021147196D9D339 CRC64;

Query Match
Best Local Similarity 59.3%; Score 101; DB 11; Length 472;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 3 RSGSFEVGLFGQHPDYFLMEPAMHV 29
DB 136 RTGSSFEVGFENNDIFLFEPLMHI 162
|||||
ID 090212 PRELIMINARY; PRT; 474 AA.
AC 090212;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN Rattus norvegicus (Rat).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Sciurognathi; Muridae; Murinae; Rattus.
OX Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT "Chondroitin 6-sulfotransferase of rat sciatic nerve."
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Li X., Kwok C.-F., Shum K.-Y.;
RT Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=21-DAY POST-CRUSH SCIATIC NERVE;
RA Liu J., Li X., Kwok C.-F., Tang S., Shum K.-Y.;
RT Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF178689; AAD54386.2;

```

KW Transferrase.
SQ SEQUENCE 474 AA; 54018 MW; 83EC94C3894CE77E CRC64;

Query Match 60.8%; Score 101; DB 11; Length 474;
Best Local Similarity 59.3%; Pred. No. 5e-06;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSFFVGLGCGHPDVFILMEPAHV 29
138 RTGSSFFVGEFFNQGNIFLFEPLMHI 164

RESULT 20

075099 PRELIMINARY; PRT; 479 AA.

AC 075099; PRELIMINARY; PRT; 479 AA.
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE Chondroitin 6-sulfotransferase (EC 2.8.2.17) (C69T).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RP TISSUE=FETAL BRAIN;
RC MEDLINE=98390270; PubMed=9714735;
RA Fukuta M., Kobayashi Y., Uchinura K., Kimata K., Hahuchi O.;
RT "Molecular cloning and expression of human chondroitin 6-sulfotransferase.";
RL Biochim. Biophys. Acta 1399:57-61(1998).
[2]
RN SEQUENCE FROM N.A.
RP TISSUE=PLACENTA;
RC MEDLINE=99098360; PubMed=9883891;
RA Tsutsuni K., Shimakawa H., Kitagawa H., Sugahara K.;
RT "Functional expression and genomic structure of human chondroitin 6-sulfotransferase.";
RL FEBS Lett. 441:235-241(1998).
CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN SULFATE.
CC -1- CATALYTIC ACTIVITY: 3'-PHOSPHOADENYLYLSULFATE + CHONDROITIN = ADENOSINE 3',5'-BISPHOSPHATE + CHONDROITIN 6'-SULFATE.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE (BY SIMILARITY).
CC -1- TISSUE SPECIFICITY: FOUND AT HIGHEST LEVELS IN SKELETAL MUSCLE.
CC ALSO PRESENT AT LOWER LEVELS IN HEART, PLACENTA AND PANCREAS.
DR EMBL: AB012192; BAA32576.1; -
DR EMBL: AB017915; BAA36348.1; -
KM Transferrase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 24 38 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (BY SIMILARITY).
FT DOMAIN 39 458 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 420 420 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 464 464 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 387 387 P -> R (IN PLACENTA).
FT VARIANT 443 443 P -> A (IN PLACENTA).
SQ SEQUENCE 479 AA; 54672 MW; 494A58A48E37D570 CRC64;

Query Match 60.8%; Score 101; DB 4; Length 479;
Best Local Similarity 59.3%; Pred. No. 5e-06;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSFFVGLGCGHPDVFILMEPAHV 29
142 RTGSSFFVGEFFNQGNIFLFEPLMHI 168

RESULT 21
09VMC3 PRELIMINARY; PRT; 315 AA.

AC 09VMC3; PRELIMINARY; PRT; 315 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE CG9551 protein.
DN CG9551.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=BERKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Aghayani A., An H.-J., Andrews-Frankoch C., Baldwin D.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu I., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Gloder A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kension J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Mostreli A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson K., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Teclor C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of Drosophila melanogaster.";
DR Science 287:2185-2195(2000).
DR EMBL: AB003614; AAF52398.1; -
DR Flybase; FBgn0031827; CG9551.
SQ SEQUENCE 315 AA; 35608 MW; EA0A82855CC6418 CRC64;

Query Match 42.8%; Score 71; DB 5; Length 315;
Best Local Similarity 52.0%; Pred. No. 0.054;
Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 SRRGSSFFVGLGCGHPDVFILMEP 25
Db 191 SRRGSSFFVGLGCGHPDVFILMEP 215

RESULT 22
095TN9 PRELIMINARY; PRT; 486 AA.

```

AC Q95YN9;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE LD30505IP.
GN CG9551.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
PT Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RA Stalton M., Brooks P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nuno J., Pacled J., Paragas V., Park S., Prounenavong S., Wan K.,
RA Yu C., Lewis S.E., Rudin G.M., Gelinkov S.,
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY058647; AAL13876.1;
DR Flybase; FBgn0031827; CG9551.
DR InterPro; IPR000683; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase; 1.
SQ SEQUENCE 486 AA; 55539 MW; BECC96F096746EB CRC64;

Query Match 42.8%; Score 71; DB 5; Length 486;
Best Local Similarity 52.0%; Pred. No. 0.087;
Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 SWSSSFVGLFGQHPDVFYLMPEP 25
DB 191 SWRSGSTFLGDIUNSLPGNFHYEP 215

RESULT 23
Q8TG14 PRELIMINARY; PRT; 46 AA.
AC Q8TG14;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mating type 1 protein (Fragment).
OS Phaeosphaeria avenaria.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes Incertae sedis;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=55068;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 58582;
RA Czembor P.C., Arseniuk E., Bergstrom G.C., Teng P.P.;
RT "Sequence diversity of mating type genes in Phaeosphaeria avenaria.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY072936; AAL69556.1;
FT NON_TER 1
FT NON_TER 46
FT SEQUENCE 46 AA; 5554 MW; 48A28037B52FF32 CRC64;

Query Match 34.3%; Score 57; DB 3; Length 46;
Best Local Similarity 50.0%; Pred. No. 0.62;
Matches 12; Conservative 5; Mismatches 5; Indels 2; Gaps 1;

QY 6 SSFVGLFGQHPD--VEYLMEPAM 27
DB 20 SSYIGQLMKQEPDKPLMTITKAM 43

RESULT 24
Q8YWG4 PRELIMINARY; PRT; 1570 AA.
AC Q8YWG4;
DT 01-MAR-2002 (TREMBLrel. 20, Created)

```

```

DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein A11649.
GN A11649.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimizu S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003586; BAB78015.1;
DR InterPro; IPR001227; Acyltransferase.
DR InterPro; IPR000794; Ketoacyl-synth.
DR InterPro; IPR003880; Pantone-attach.
DR Pfam; PF00698; Acyl-transferase; 1.
DR Pfam; PF00109; ketoacyl-synth; 1.
DR Pfam; PF02801; ketoacyl-synth; 1.
DR Pfam; PF00550; PP-binding; 1.
DR PROSITE; PS0075; ACP-DOMAIN; 1.
DR PROSITE; PS00606; B_KETOACYL-SYNTHASE; 1.
DR Hypothetical protein; Complete proteome.
SQ SEQUENCE 1570 AA; 170791 MW; AB778DE7395F6A52 CRC64;

Query Match 34.3%; Score 57; DB 16; Length 1570;
Best Local Similarity 48.5%; Pred. No. 30;
Matches 16; Conservative 2; Mismatches 7; Indels 8; Gaps 2;

QY 4 SGSS----FVGLFGQHPDVFYLMPEP--AMW 28
DB 692 SGQSQAQIFVTQTQQGHPDIFKELPVNKAH 724

RESULT 25
Q99VX8 PRELIMINARY; PRT; 277 AA.
AC Q99VX8;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Telchoic acid translocation permease protein.
GN TAGG OR SA0594.
OS Staphylococcus aureus (strain N315).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC Staphylococcus.
OX NCBI_TaxID=158879;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Ito T., Kanamori M.,
RA Matsumaru H., Matuyama A., Murakami H., Hosoyama A., Mizutani-Ui Y.,
RA Takahashi N.K., Sawano T., Inoue R.T., Kaito C., Sekimizu K.,
RA Hirakawa H., Kuhara S., Goto S., Yabuzaki J., Kanehisa M.,
RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,
RA Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant staphylococcus
aureus."
RL Lancet 357:1225-1240(2001).
DR EMBL; AP003137; BAB41826.1;
DR InterPro; IPR00412; ABCtransporter2.
DR Pfam; PF01061; ABC2_membrane; 1.
DR Complete proteome.
SQ SEQUENCE 277 AA; 32381 MW; 603C9629F436608 CRC64;

Query Match 31.9%; Score 53; DB 16; Length 277;

```


DR EMBL: AE009866; AAL64063.1; -
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 187 AA; 21760 MW; 0A33A9408FE41942 CRC64;

Query Match 31.6%; Score 52.5; DB 17; Length 187;
 Best Local Similarity 37.0%; Pred. No. 12;
 Matches 10; Conservative 6; Mismatches 6; Indels 5; Gaps 1;

QY 2 RRGSSFGQLFGQHPD-----PDVEYLM 23
 ID 09C409 PRELIMINARY; PRT; 46 AA.
 AC 09C409;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE Mating type 1 protein (Fragment).
 OS Phaeosphaeria nodorum (Septoria nodorum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
 OC Phaeosphaeriaceae; Phaeosphaeria.
 NCBI_TaxID=13684;
 RN (1)
 RC STRAIN=SN48-1;
 RP SEQUENCE FROM N.A.
 RA Dai O., Arseniuk R., Cui K., Deng P.P.;
 RT "Genetic segregation and sexuality in Phaeosphaeria nodorum.";
 RL Submitted (NOV-2000) to the EMBL/Genbank/DBD databases.
 DR EMBL: AF322008; AAK09389.1; -
 FT NON_TER 1
 SO SEQUENCE 46 AA; 5553 MW; 48A678A4439DB4D2 CRC64;

Query Match 31.3%; Score 52; DB 3; Length 46;
 Best Local Similarity 45.8%; Pred. No. 3.1;
 Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 6 SSFVGQLFGQHPD--VFYLMPEAW 27
 ID 08TG16 PRELIMINARY; PRT; 46 AA.
 AC 08TG16;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Mating type 1 protein (Fragment).
 OS Phaeosphaeria nodorum (Septoria nodorum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
 OC Phaeosphaeriaceae; Phaeosphaeria.
 NCBI_TaxID=13684;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Dai O., Cui K.-R., Arseniuk E., Deng P.P.;
 RT "Segregation, Aggressiveness, and Sexuality in Phaeosphaeria
 RT nodorum.";
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBD databases.
 DR EMBL: AT072933; AAL69553.1; -
 FT NON_TER 1
 SO SEQUENCE 46 AA; 5614 MW; 48A678A45E2B6FA2 CRC64;

Query Match 31.3%; Score 52; DB 3; Length 46;
 Best Local Similarity 45.8%; Pred. No. 3.1;

Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 6 SSFVGQLFGQHPD--VFYLMPEAW 27
 ID 08TG16 PRELIMINARY; PRT; 46 AA.
 AC 08TG16;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Mating type 1 protein (Fragment).
 OS Phaeosphaeria nodorum (Septoria nodorum).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
 OC Phaeosphaeriaceae; Phaeosphaeria.
 NCBI_TaxID=13684;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA Dai O., Cui K.-R., Arseniuk E., Deng P.P.;
 RT "Segregation, Aggressiveness, and Sexuality in Phaeosphaeria
 RT nodorum.";
 RL Submitted (JAN-2002) to the EMBL/Genbank/DBD databases.
 DR EMBL: AT072933; AAL69553.1; -
 FT NON_TER 1
 SO SEQUENCE 46 AA; 5614 MW; 48A678A45E2B6FA2 CRC64;

Query Match 31.3%; Score 52; DB 16; Length 518;
 Best Local Similarity 47.8%; Pred. No. 44;
 Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 3 RRGSSFGQLFGQHPDVFYLMPE 25
 ID 08U1V4 PRELIMINARY; PRT; 614 AA.
 AC 08U1V4;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE ABC transporter, substrate binding protein.
 GN ATU0187 OR AGR_C_316.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 NCBI_TaxID=176299;
 RN (1)
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21608550; Pubmed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F. Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
 RA Chapman P., Clendenning U., Deatherage G., Gillet W., Grant C.,
 RA Kutayavitt T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
 RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Nester E.W.;
 RT "The genome of the natural genetic engineer Agrobacterium tumefaciens

Query Match 31.3%; Score 52; DB 16; Length 518;
 Best Local Similarity 47.8%; Pred. No. 44;
 Matches 11; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

```

RT CS8."
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA Ouyollo B., Goldman B.S., Cao Y., Askenzi M., Halling C., Mullin L.,
RA Houniel K., Gordon J., Vaudin M., Iartchouk O., Bpp A., Liu F.,
RA William C., Allinger M., Doughly D., Scott C., Lappas C., Markelz B.,
RA Flanagan C., Crowell C., Garson J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens CS8."
RL Science 294:2323-2328(2001).
DR EMBL; AF008992; AAL4211.1;
DR EMBL; AF007959; AAK6007.1;
KW Complete proteome.
SQ SEQUENCE 614 AA; 68551 MW; 117D62469C0F3ED5 CRC64;

Query Match
Best Local Similarity 31.3%; Score 52; DB 16; Length 614;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 2 WRSQSPVQGLFGQHPDVE 20
DB 33 WRKGISTVQGL--KHDPGF 49
||| | | | | | | | | |
| | | | | | | | | |

RESULT 34
Q8Y756 PRELIMINARY; PRT; 641 AA.
AC Q8Y756;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein A112870.
GN A112870.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11759840;
RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120."
RL DNA Res. 8:205-213(2001).
DR EMBL; AP003591; BAB74569.1;
DR InterPro; IPR003342; PWT.
DR Pfam; PF02366; PWT.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 641 AA; 71650 MW; 8A709F53750AF2C CRC64;

Query Match
Best Local Similarity 31.3%; Score 52; DB 16; Length 641;
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 2 WRSQSPVQGLFGQHPDVE 16
DB 298 WRNGESTVNSFTGYH 312
||| | | | | | | | | |
| | | | | | | | | |

RESULT 35
Q96X13 PRELIMINARY; PRT; 46 AA.
AC Q96X13;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

```

```

DE Mating type 1 protein (Fragment).
OS Phaeosphaeria avenaria f. sp. triticeae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae sedis;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=54790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA125-1;
RA Dai Q., Cui K., Arseniuk E., Ueng P.P.;
RT "Mating type gene in Phaeosphaeria avenaria f. sp. triticeae."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF558824; AAK51440.1;
FT NON_TER 1 46
FT NON_TER 1 46
SQ SEQUENCE 46 AA; 5645 MW; 48BCB74461858F4C CRC64;

Query Match
Best Local Similarity 30.7%; Score 51; DB 3; Length 46;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 6 SSFVQGLFGQHPD--VFYLMEPAM 27
DB 20 SSYIGQLMKYEPDKPLMTLTKAM 43
||:|:|:| | | | | | | |
| | | | | | | | | |

RESULT 36
Q96X15 PRELIMINARY; PRT; 46 AA.
AC Q96X15;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Mating type 1 protein (Fragment).
GN MAT1.
OS Phaeosphaeria avenaria f. sp. triticeae.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Dothideomycetes et Chaetothyriomycetes incertae Phaeosphaeriaceae;
OC Phaeosphaeriaceae; Phaeosphaeria.
OX NCBI_TaxID=54790;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SA1004CND-96;
RA Dai Q., Cui K., Gilbert J., Bergstrom G., Ueng P.P.;
RT "Mating type gene in Phaeosphaeria avenaria f. sp. triticeae."
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 18596;
RA Czenbor P.C., Arseniuk E., Ueng P.P.;
RT "Sequence diversity of mating type genes in Phaeosphaeria avenaria f. sp. triticeae."
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF072935; AAL68555.1;
DR EMBL; AF077724; AAL82533.1;
FT NON_TER 1 46
FT NON_TER 1 46
SQ SEQUENCE 46 AA; 5644 MW; 48BCB74461858F4C CRC64;

Query Match
Best Local Similarity 30.7%; Score 51; DB 3; Length 46;
Matches 11; Conservative 5; Mismatches 6; Indels 2; Gaps 1;

QY 6 SSFVQGLFGQHPD--VFYLMEPAM 27
DB 20 SSYIGQLMKYEPDKPLMTLTKAM 43
||:|:|:| | | | | | | |
| | | | | | | | | |

```

```

RESULT 37
Q9XBL3      PRELIMINARY;      PRT;      237 AA.
ID  Q9XBL3;
AC  Q9XBL3;
DT  01-NOV-1999 (TREMBLrel. 12, Created)
DT  01-NOV-1999 (TREMBLrel. 12, last sequence update)
DE  01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE  DNA alkylation repair enzyme.
GN  ALKD.
OS  Bacillus cereus.
OC  Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
OC  Bacillaceae; Bacillus.
OX  NCBI_TaxID=1396;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 10987;
RX  MEDLINE=99231848; PubMed=10217496;
RA  Okstad O.A., Hegna I., Lindaeck T., Rishovd A.L., Kolsto A.B.;
RT  "Genome organisation is not conserved between Bacillus cereus and
RL  Bacillus subtilis.";
RL  Microbiology 145:621-631(1999).
DR  EMBL: AJ010128; CAB40581.1;
SQ  SEQUENCE 237 AA; 28172 MW; 132DAF8961496A83 CRC64;

Query Match      30.7%; Score 51; DB 2; Length 237;
Best Local Similarity 38.1%; Pred. No. 26;
Matches 8; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

QY  7 SFVGOLFQHPDVFYLMPEAW 27
DB  118 TFLGNIFLQHPDLISATPKW 138

RESULT 38
Q9LJS3      PRELIMINARY;      PRT;      263 AA.
ID  Q9LJS3;
AC  Q9LJS3;
DT  01-OCT-2000 (TREMBLrel. 15, Created)
DT  01-OCT-2000 (TREMBLrel. 15, last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE  Putative ATP-binding subunit.
GN  URTD.
OS  Anabaena sp. (strain PCC 7120).
OC  Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX  NCBI_TaxID=103690;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=PCC 7120;
RA  Valladares A., Montesinos M., Herrero A., Flores F.;
RT  "A procarboxylic urea transporter identified in two cyanobacteria.";
RL  Submitted (JAN-2000) to the EMBL/GenBank/DBD databases.
DR  EMBL: AA271599; CAB70951.1;
DR  HSSP: Q58663; 1G6H;
DR  InterPro: IPR003439; ABC_transport.
DR  Pfam: PF00005; ABC_tran; 1.
KW  ATP-binding.
SQ  SEQUENCE 263 AA; 30320 MW; F3E09CCC95B6A8 CRC64;

Query Match      30.7%; Score 51; DB 2; Length 263;
Best Local Similarity 40.0%; Pred. No. 29;
Matches 10; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY  1 SWRSGSFVGOLFQHPDVFYLMPE 25
DB  166 SWRKATFRMKLVGQSPDLVDEP 190

RESULT 39
Q93WX7      PRELIMINARY;      PRT;      341 AA.
AC  Q93WX7;

```

```

DT  01-DEC-2001 (TREMBLrel. 19, Created)
DT  01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT  01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE  Aspartate aminotransferase (Fragment).
GN  CRAT1.
OS  Securigera parviflora.
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids I; Fabales; Fabaceae; Papilionoideae; Loreae; Securigera.
OX  NCBI_TaxID=172083;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=LEAF-DERIVED CALLUS;
RA  Arlanti N.;
RT  "Molecular characterization of selected genes in ammonium assimilation
RL  from in vitro cultures of the legume Coronilla rostrata.";
RL  Thesis (2001). The University of New South Wales.
DR  EMBL: AF419301; AAL09704.1;
DR  InterPro: IPR004839; AminoTransf1/2.
DR  InterPro: IPR000796; AsptTransf_sub.
DR  InterPro: IPR004838; Nitransf_1.
DR  Pfam: PF00155; aminotran_1.2; 1.
DR  PRINTS: PR00799; TRANSAMINASE.
DR  PROSITE: PS00105; AA_TRANSFER_CLASS_1; UNKNOWN_1.
KW  Aminotransferase; Transferase.
FT  NON_TER
SQ  SEQUENCE 341 AA; 37396 MW; 915AB17531FD092E CRC64;

Query Match      30.7%; Score 51; DB 10; Length 341;
Best Local Similarity 37.0%; Pred. No. 39;
Matches 10; Conservative 5; Mismatches 10; Indels 2; Gaps 1;

QY  1 SWRSGSFVGOLFQHPDVFYLMPEAW 27
DB  46 SLRVGGEPLGKHY--HQRIVLPQPTW 70

RESULT 40
Q8VP21      PRELIMINARY;      PRT;      681 AA.
ID  Q8VP21;
AC  Q8VP21;
DT  01-MAR-2002 (TREMBLrel. 20, Created)
DT  01-MAR-2002 (TREMBLrel. 20, last sequence update)
DE  01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE  Bscsc.
OS  Burkholderia cepacia genomovar III.
OC  Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC  Burkholderia.
OX  NCBI_TaxID=95486;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  MEDLINE=21441674; PubMed=11557147;
RA  Parsons T.N., Glendinning K.J., Thornton V., Hales B.A., Hart C.A.,
RA  Winstanley C.;
RT  "A putative type III secretion gene cluster is widely distributed in
RL  the Burkholderia cepacia complex but absent from genomovar I.";
RL  FEMS Microbiol. Lett. 203:103-108(2001).
DR  EMBL: AY028431; AAL59397.1;
DR  InterPro: IPR004845; GSP11/IIIProtein.
DR  InterPro: IPR004845; GSP11/IIIProtein.
DR  Pfam: PF00263; GSP11-II; 1.
DR  PROSITE: PS00875; TSP_D; UNKNOWN_1.
SQ  SEQUENCE 681 AA; 72502 MW; 9C512C0F8A05AE0B CRC64;

Query Match      30.7%; Score 51; DB 2; Length 681;
Best Local Similarity 36.7%; Pred. No. 63;
Matches 11; Conservative 3; Mismatches 10; Indels 6; Gaps 1;

QY  2 WRSGSFVGOLF-----GHPDVFYLMPE 25
DB  568 WLSDIPLIGLNFKYTDKSGNHMERFYLLTP 597

```

```

RESULT 41
OQ25D2 PRELIMINARY: PRT: 158 AA.
AC Q925D2:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE Bcl2-interacting killer-like protein.
GN BIKLK.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY;
RA Itoh T., Itoh A., Pleasure D.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF372501; AAK53820.1; -
SQ SEQUENCE 158 AA; 17660 MW; A63742550879C78 CRC64;

Query Match
Best Local Similarity 41.4%; Pred. No. 23;
Matches 12; Conservative 3; Mismatches 14; Indels 0; Gaps 0;

OQ 1 SMRSGSFVGLFGQHPDYFLMEPAHV 29
Db 127 AMVSPDQDBQLPQVYLVLGLGAMHL 155

RESULT 42
OQVNG5 PRELIMINARY: PRT: 239 AA.
AC OQVNG5:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Hypothetical 26.7 kDa protein.
OS Pseudomonas putida.
OC Pseudomonas.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX NCBI_TaxID=303;
RN [1]
RP SEQUENCE FROM N.A.
RA Greated A., Lambertson L., Williams P.A., Thomas C.M.;
RT "Complete nucleotide sequence of IncP-9 plasmid PMW0."
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ344068; CAC86851.1; -
KM Hypothetical protein; Plasmid.
SQ SEQUENCE 239 AA; 26726 MW; A65732EA71C5B0A6 CRC64;

Query Match
Best Local Similarity 47.4%; Pred. No. 36;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OQ 8 FVGLFGQHPDYFLMEPA 26
Db 83 FMGQZGSHEDIDYFVPA 101

RESULT 43
OQJN46 PRELIMINARY: PRT: 678 AA.
AC OQJN46:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative alpha amylase (Fragment).
GN AAM1.
OS Rhodospirillum rubrum (Rhodospirillum rubrum).
OC Bacteria; Proteobacteria; alpha subdivision; Rhodospirillum rubrum;
OC Rhodospirillum rubrum.

```

```

OX NCBI_TaxID=1063;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=96079285; PubMed=7494484;
RA Ward M.J., Bell A.W., Hamblin P.A., Packer H.L., Armitage J.P.;
RT "Identification of a chemotaxis operon with two che genes in
RT Rhodospirillum rubrum."
RL Mol. Microbiol. 17:357-366(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=96154945; PubMed=8596451;
RA Ward M.J., Harrison D.M., Eber M.J., Armitage J.P.;
RT "Identification of a methyl-accepting chemotaxis protein in
RT Rhodospirillum rubrum."
RL Mol. Microbiol. 18:115-121(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=MS8N;
RX MEDLINE=20428429; PubMed=10970853;
RA Shah D.S.H., Porter S.L., Martin A.C., Hamblin P.A., Armitage J.P.;
RT "Fine tuning bacterial chemotaxis: Analysis of Rhodospirillum rubrum
RT behaviour under aerobic and anaerobic conditions by mutation of the
RT major chemotaxis operons and che genes."
RL EMBL: J. 19:4601-4613(2000).
DR EMBL: X80205; CAB87126.1; -
DR InterPro: IPR000461; Alpha_Amylase.
DR InterPro: IPR000294; Vtk_dep_GLA.
DR Pfam: PF00128; alpha-amylase; 1.
DR PROSITE: PS00011; GIU-CARBOXYLATION; UNKNOWN_1.
FT NON TER 678
SQ SEQUENCE 678 AA; 77092 MW; E88BDD0339F1B61D CRC64;

Query Match
Best Local Similarity 50.0%; Pred. No. 11e+02;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OQ 9 VGLFGQHPDYFLME 24
Db 475 IGEVQSGHPDYFLAE 490

RESULT 44
OQJN4 PRELIMINARY: PRT: 1051 AA.
AC OQJN4:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE WD repeat domain protein.
GN F14L2-80.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=5702;
RN [1]
RP SEQUENCE FROM N.A.
RA Jordan N., Bangert S., Wiedelmann R., Voss H., Unsel M., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X., Quetier F., Salanoubat M.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA ED Arabidopsis sequencing project;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC - SIMILARITY: CONTAINS 6 WD REPEATS (TRP-ASP DOMAINS).
DR EMBL: AL353818; CAB88535.1; -
DR InterPro: IPR001680; WD40.
DR Pfam: PF00400; WD40; 7.
DR PRINTS: PR00320; GPROTEINRPT.
DR ProDom: PD000018; WD40; 2.
DR SMART: SM00320; WD40; 5.

```

DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 DR PROSITE; PS50082; WD_REPEATS_2; 3.
 DR PROSITE; PS50294; WD_REPEATS_REGION; 1.
 KW Repeat; WD repeat.
 SQ SEQUENCE 1051 AA; 115562 MW; 881P9E2A826778C9 CRC64;

Query Match
 Best Local Similarity 30.1%; Score 50; DB 10; Length 1051;
 Matches 11; Conservative 5; Mismatches 12; Indels 2; Gaps 1;

QY 1 SMRSGSFYQGLFGQHPDV--FYLMPEAWH 28
 Db 823 SMNIGSAFCGLAGLDYDKYMARKGN 852

RESULT 45
 ID 099140 PRELIMINARY; PRT; 1645 AA.
 AC 099140;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE Beta-H spectrin (Fragment).
 GN KST OR CG12008.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN RP
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEAD;
 RX MEDLINE=91035599; PubMed=2229176;
 RA Dubreuil R.R., Byers T.J., Stewart C.T., Kiehart D.P.;
 RT "A beta-spectrin isoform from Drosophila (beta H) is similar in size
 RT to vertebrate dystrophin.";
 RL J. Cell Biol. 111:1849-1858(1990).
 CC -1- FUNCTION: SPECTRIN IS A MAJOR CONSTITUENT OF THE CYTOSKELETON
 CC NETWORK. THE NATIVE SPECTRIN FORMED FROM THE BETA-H SUBUNIT,
 CC RESEMBLES A NON-EYTHROID TYPE.
 CC -1- FUNCTION: BETA SPECTRIN ALONE MAY BE IMPORTANT FOR RECEPTOR
 CC CLUSTERING.
 CC -1- SUBUNIT: NATIVE SPECTRIN MOLECULE IS A TETRAMER COMPOSED OF TWO
 CC ANTIPARALLEL HETERODIMERS JOINED HEAD TO HEAD SO THAT EACH END OF
 CC THE NATIVE MOLECULE INCLUDES THE C-TERMINUS OF THE ALPHA SUBUNIT
 CC AND THE N-TERMINUS OF THE BETA-H SUBUNIT.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 DR EMBL X53997; CAA37939.1; -.
 DR PIR; A37792; A37792.
 DR PIR; S15666; S15666.
 DR HSSP; Q01082; 1BKR.
 DR FLYBASE; F8an0004167; kst.
 DR InterPro; IPR001589; Actbind-actinh.
 DR InterPro; IPR001715; Calponin-like.
 DR InterPro; IPR001452; SH3.
 DR InterPro; IPR002017; Spectrin.
 DR Pfam; PF00307; CH; 2.
 DR Pfam; PF00435; Spectrin; 1.
 DR Pfam; PF00435; Spectrin; 13.
 DR ProDom; PD000066; SH3; 1.
 DR SMART; SM00033; CH; 2.
 DR SMART; SM00326; SH3; 1.
 DR SMART; SM00150; SPEC; 10.
 DR PROSITE; PS00019; ACTININ_1; 1.
 DR PROSITE; PS00020; ACTININ_2; 1.
 DR PROSITE; PS00021; CH; 2.
 DR PROSITE; PS00002; SH3; 1.
 KM Actin-binding; Capping protein; Cytoskeleton; Membrane; Repeat;
 KM SH3 domain.
 FT NON_TER 1 1
 FT DOMAIN 1 259 SEGMENT 1.
 FT DOMAIN 260 371 SEGMENT 2.
 FT DOMAIN 372 482 SEGMENT 3.

FT DOMAIN 801 961 SEGMENT 7.
 FT REPEAT 483 588 SEGMENT 4.
 FT REPEAT 589 694 SEGMENT 5.
 FT REPEAT 695 800 SEGMENT 6.
 FT REPEAT 962 1063 SEGMENT 8.
 FT REPEAT 1064 1169 SEGMENT 9.
 FT REPEAT 1170 1274 SEGMENT 10.
 FT REPEAT 1275 1380 SEGMENT 11.
 FT REPEAT 1381 1484 SEGMENT 12.
 FT REPEAT 1485 1590 SEGMENT 13.
 FT REPEAT 1591 1645 SEGMENT 14.
 FT NON_TER 1645 1645
 SQ SEQUENCE 1645 AA; 190480 MW; 22FEDCBFB83023EF CRC64;

Query Match
 Best Local Similarity 45.0%; Score 50; DB 5; Length 1645;
 Matches 9; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

QY 1 SMRSGSFYQGLFGQHPDV 20
 Db 186 SMRSGFGNALHSHRPDLF 205

RESULT 46
 ID 09VZ03 PRELIMINARY; PRT; 4097 AA.
 AC 09VZ03; 044084;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE KST protein (Beta-H spectrin).
 GN KST OR KARST OR CG12008.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H.C., Blazek R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abiri J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bertan B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Mervulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pauley J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Ventier E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,

```

TR      opportunistic pathogen." ;
RL      Nature 406:959-964(2000) ;
DR      EMBL: AE004525; AAC04306.1; -.
DR      InterPro: IPR003855; K+-transprtr.
DR      Pfam: PF02705; K-trans; 1.
KW      Complete proteome.
SQ      SSOURCE      634 AA; 68734 MW; 504CBA5F70CCBE99 CRC64;

Query Match      29.8%; Score 49.5; DB 16; Length 634;
Best Local Similarity 26.1%; Pred. No. 1,2e+02;
Matches 12; Conservative 5; Mismatches 10; Indels 19; Gaps 1.

OY      3 RSGGSFVGOLFEG-----QHDFVFLMEPPAHV 29
          : : : : : : : : : : : : : : : : : :
DB      171 KHGTARIGILFGFVYLMFGALALAGVYGVYIQDPEVLQANPVMVAV 216

RESULT 48
ID      Q9NV23      PRELIMINARY;      PRT;      265 AA.
AC      Q9NV23;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE      CNNA FL11092.fis, clone PLACL001692, moderately similar to S-acyl
DE      fatty acid synthase thioesterase, medium chain (EC 3.1.2.14) .
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RA      Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA      Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,
RA      Wakamatsu A., Nakamura Y., Nagahari K., Masuko Y., Oshima A.;
RT      "MDO human cDNA sequencing project.";
RL      Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL: AK001844; BAA91937.1; -.
DR      InterPro: IPR000379; Ser_estrs_site.
DR      InterPro: IPR001031; Thioesterase.
DR      Pfam: PF00975; Thioesterase; 1.
SQ      SEQUENCE      265 AA; 29931 MW; C67722F3CAD2D2C7 CRC64;

Query Match      29.5%; Score 49; DB 4; Length 265;
Best Local Similarity 41.4%; Pred. No. 56;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2.

OY      1 SMR---SGGSFVGOLFQHPDFVFLMEPPA 26
          : : : : : : : : : : : : : : : : : :
DB      219 AKMDVTSGNAKIYQLPGH---FYLLDPA 244

RESULT 49
ID      Q9NTM1      PRELIMINARY;      PRT;      318 AA.
AC      Q9NTM1;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE      CNNA FL11106.fis, clone PLACL005763, moderately similar to S-acyl
DE      fatty acid synthase thioesterase, medium chain (EC 3.1.2.14) .
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX      NCBI_Taxid=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      TISSUE=PLACENTA;
RA      Iisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
RA      Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,

```

RA Matsunawa H., Ishii S., Kawai Y., Saigo K., Yamamoto J., Wakamatsu A.,
 RA Nakamura Y., Nagahara K., Masuno Y., Sasaki N.,
 RT "NEDO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK001968; BAA92007.1; -
 DR InterPro: IPR001031; Thioesterase.
 DR Pfam: PF00975; Thioesterase; 1
 SQ SEQUENCE 318 AA; 35818 MW; C7DB0DEBCEB0823F CRC64;

Query Match 29.5%; Score 49; DB 4; Length 318;
 Best Local Similarity 41.4%; Pred. No. 69;
 Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR--SGSSFFVGLFGQHPDVFLMEPA 26
 Db 272 AKKDVTSGNKRYQLPFGH--FYLLDPA 297

RESULT 50

095UT2 PRELIMINARY; PRT; 891 AA.
 ID 095UT2;
 AC Q95UT2;
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)
 DT 01-DEC-2001 (TrEMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, last annotation update)
 DE Iron regulatory protein 1.
 OS Manduca sexta (Tobacco hawkmoth) (Tobacco hornworm).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 OC Spingidae; Spingidae; Spinginae; Manduca.
 OX NCBI_TaxID=7130;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Zhang D., Ferris C., Gailer J., Kohlhepp P., Winzerling J.J.;
 RT "Manduca sexta IRPI: molecular characterization and in vivo response
 to iron.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY032658; AAK39637.1; -
 DR InterPro: IPR000573; Aconitase_C.
 DR InterPro: IPR001030; Aconitase_N.
 DR Pfam: PF00330; Aconitase; 1.
 DR Pfam: PF00694; Aconitase; 1.
 DR ProDom: PD000511; Aconitase_N; 1.
 DR ProSITE; PS00450; ACONITASE_1; UNKNOWN_1.
 DR ProSITE; PS01244; ACONITASE_2; UNKNOWN_1.
 SQ SEQUENCE 891 AA; 97268 MW; 592D36D150E041E4 CRC64;

Query Match 29.5%; Score 49; DB 5; Length 891;
 Best Local Similarity 40.0%; Pred. No. 2.1e+02;
 Matches 10; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 SWRSGSFFVGLFGQHPDVFLMEP 25
 Db 626 SWQSLVPGKLYGWDNSTRYIKKP 650

Search completed: February 20, 2003, 13:33:24
 Job time : 73.4286 secs

This page Blank (uspio)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:31 ; Search time 33.1429 Seconds
(without alignments)
25.745 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166
Sequence: 1 SWMSGSFVGOLEFGHPDYIMEPMNV 29

Scoring table: BLOSUM62
Gap: 10.0, Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 1000 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	166	100.0	US-09-045-284A-2	Sequence 2, Appli
2	166	100.0	US-09-190-911-1	Sequence 1, Appli
3	124	74.7	US-09-263-023-2	Sequence 2, Appli
4	124	74.7	US-09-471-867-2	Sequence 2, Appli
5	124	74.7	US-09-263-023-4	Sequence 4, Appli
6	124	74.7	US-09-471-867-4	Sequence 4, Appli
7	116	69.9	US-09-015-188-2	Sequence 2, Appli
8	101	60.8	US-08-655-878-2	Sequence 2, Appli
9	101	60.8	US-08-899-514-2	Sequence 2, Appli
10	53.5	32.2	US-08-443-865-6	Sequence 6, Appli
11	53.5	32.2	US-08-900-491-6	Sequence 6, Appli
12	47	28.3	US-09-202-491-7	Sequence 7, Appli
13	46.5	28.0	US-08-737-825-6	Sequence 6, Appli
14	45	27.1	US-09-134-001C-5331	Sequence 5331, Ap
15	45	27.1	US-08-551-386-4	Sequence 4, Appli
16	45	27.1	PCT-US93-12687-4	Sequence 4, Appli
17	45	27.1	US-08-206-176-2	Sequence 2, Appli
18	45	27.1	US-09-643-597-170	Sequence 170, App
19	45	27.1	US-09-373-157-4	Sequence 4, Appli
20	45	27.1	US-09-643-597-357	Sequence 357, App
21	45	27.1	US-09-193-562D-32	Sequence 32, Appli
22	45	27.1	US-09-643-597-161	Sequence 161, App
23	45	27.1	US-08-551-386-6	Sequence 6, Appli
24	45	27.1	PCT-US93-12687-6	Sequence 6, Appli
25	44	26.5	US-09-357-251-22	Sequence 22, Appli
26	44	26.5	US-09-052-877-2	Sequence 2, Appli
27	44	26.5	US-08-740-223A-15	Sequence 15, Appli

28	44	26.5	496	4	US-09-709-188-15	Sequence 15, Appli
29	44	26.5	627	1	US-08-295-814E-4	Sequence 4, Appli
30	44	26.5	627	1	US-08-291-299-10	Sequence 10, Appli
31	44	26.5	627	4	US-09-343-361-4	Sequence 4, Appli
32	44	26.5	627	5	PCT-US93-01959-14	Sequence 14, Appli
33	44	26.5	627	5	PCT-US93-01959-10	Sequence 10, Appli
34	44	26.5	632	1	US-08-295-814E-10	Sequence 10, Appli
35	44	26.5	632	1	US-09-343-361-10	Sequence 10, Appli
36	44	26.5	632	5	PCT-US93-01959-10	Sequence 10, Appli
37	44	26.5	1271	1	US-08-095-734-2	Sequence 2, Appli
38	44	26.5	1271	2	US-08-444-623-2	Sequence 2, Appli
39	44	26.5	1271	3	US-08-471-869-2	Sequence 2, Appli
40	44	26.5	1271	5	US-09-342-563-2	Sequence 2, Appli
41	44	26.5	1271	5	PCT-US94-08267-2	Sequence 2, Appli
42	44	26.5	2509	4	US-08-469-005A-10	Sequence 10, Appli
43	44	26.5	2511	4	US-09-261-907-2	Sequence 2, Appli
44	44	26.5	256	3	US-08-906-797-89	Sequence 89, Appli
45	44	26.5	256	3	US-08-906-616-89	Sequence 89, Appli
46	44	26.5	256	4	US-08-817-795-89	Sequence 89, Appli
47	44	26.5	256	4	US-08-639-075A-89	Sequence 89, Appli
48	44	26.5	256	4	US-09-012-431-89	Sequence 89, Appli
49	44	26.5	256	4	US-09-012-692-89	Sequence 89, Appli
50	44	26.5	256	4	US-08-906-613-89	Sequence 89, Appli
51	44	26.5	256	5	PCT-US93-1442A-88	Sequence 89, Appli
52	44	26.5	467	4	US-09-134-001C-4200	Sequence 4200, Ap
53	44	26.5	22	1	US-08-003-966-3	Sequence 3, Appli
54	44	26.5	235	1	US-08-776-059-39	Sequence 39, Appli
55	44	26.5	250	2	US-08-872-961A-1	Sequence 1, Appli
56	44	26.5	250	3	US-09-231-258-1	Sequence 1, Appli
57	44	26.5	253	4	US-08-776-059-31	Sequence 31, Appli
58	44	26.5	503	4	US-09-357-251-36	Sequence 36, Appli
59	44	26.5	504	4	US-08-776-059-35	Sequence 35, Appli
60	44	26.5	656	4	US-09-605-785-379	Sequence 379, App
61	44	26.5	656	4	US-09-439-913-379	Sequence 379, App
62	44	26.5	656	4	US-09-352-616A-379	Sequence 379, App
63	44	26.5	671	4	US-09-605-785-380	Sequence 380, App
64	44	26.5	671	4	US-09-439-913-380	Sequence 380, App
65	44	26.5	671	4	US-09-352-616A-380	Sequence 380, App
66	44	26.5	1719	4	US-09-605-785-378	Sequence 378, App
67	44	26.5	1719	4	US-09-439-913-378	Sequence 378, App
68	44	26.5	1719	4	US-09-352-616A-378	Sequence 378, App
69	44	26.5	182	4	US-09-134-001C-2842	Sequence 2842, Ap
70	44	26.5	702	4	US-09-232-200-102	Sequence 102, App
71	44	26.5	702	4	US-09-232-197-102	Sequence 102, App
72	44	26.5	702	4	US-09-232-201-102	Sequence 102, App
73	44	26.5	80	1	US-08-485-455D-55	Sequence 55, Appli
74	44	26.5	80	2	US-08-482-130C-55	Sequence 55, Appli
75	44	26.5	80	3	US-08-484-211C-55	Sequence 55, Appli
76	44	26.5	80	3	US-08-906-769-55	Sequence 55, Appli
77	44	26.5	80	3	US-08-906-616-55	Sequence 55, Appli
78	44	26.5	80	4	US-08-817-795-55	Sequence 55, Appli
79	44	26.5	80	4	US-08-485-443B-55	Sequence 55, Appli
80	44	26.5	80	4	US-08-639-075A-55	Sequence 55, Appli
81	44	26.5	80	4	US-09-012-431-55	Sequence 55, Appli
82	44	26.5	80	4	US-09-012-692-55	Sequence 55, Appli
83	44	26.5	80	4	US-08-906-613-55	Sequence 55, Appli
84	44	26.5	80	5	PCT-US93-1442A-55	Sequence 55, Appli
85	44	26.5	189	1	US-07-982-650D-2	Sequence 2, Appli
86	44	26.5	189	1	US-08-331-379-2	Sequence 2, Appli
87	44	26.5	189	1	US-07-982-650D-1	Sequence 1, Appli
88	44	26.5	189	1	US-07-982-650D-3	Sequence 3, Appli
89	44	26.5	189	1	US-07-982-650D-4	Sequence 4, Appli
90	44	26.5	189	1	US-07-982-650D-5	Sequence 5, Appli
91	44	26.5	189	1	US-08-331-379-1	Sequence 1, Appli
92	44	26.5	189	1	US-08-331-379-3	Sequence 3, Appli
93	44	26.5	189	1	US-08-331-379-4	Sequence 4, Appli
94	44	26.5	189	1	US-08-331-379-5	Sequence 5, Appli
95	44	26.5	226	4	US-09-032-515-32	Sequence 32, Appli
96	44	26.5	226	3	US-08-549-515-7	Sequence 7, Appli
97	44	26.5	349	3	US-08-465-971B-2	Sequence 2, Appli
98	44	26.5	349	3	US-08-806-597A-14	Sequence 14, Appli
99	44	26.5	349	3	US-08-970-428A-14	Sequence 14, Appli
100	44	26.5	393	4	US-08-961-083-58	Sequence 58, Appli

Line	Time	Offset	Sequence	Value
247	38	22.9	Sequence 1, Appl1	320
248	38	22.9	Sequence 27, Appl1	321
249	38	22.9	Sequence 26, Appl1	322
250	38	22.9	Sequence 25, Appl1	323
251	38	22.9	Sequence 32, Appl1	324
252	38	22.9	Sequence 36, Appl1	325
253	38	22.9	Sequence 35, Appl1	326
254	38	22.9	Sequence 2, Appl1	327
255	38	22.9	Sequence 122, App	328
256	38	22.9	Sequence 3755, App	329
257	38	22.9	Sequence 10, Appl1	330
258	38	22.9	Sequence 2, Appl1	331
259	38	22.9	Sequence 4, Appl1	332
260	38	22.9	Sequence 4054, App	333
261	38	22.9	Sequence 4, Appl1	334
262	38	22.9	Sequence 4, Appl1	335
263	38	22.9	Sequence 4, Appl1	336
264	38	22.9	Sequence 118, App	337
265	38	22.9	Sequence 6, Appl1	338
266	38	22.9	Sequence 12, Appl1	339
267	38	22.9	Sequence 18, Appl1	340
268	38	22.9	Sequence 401, App	341
269	38	22.9	Sequence 2, Appl1	342
270	37.5	22.6	Sequence 12, Appl1	343
271	37.5	22.6	Sequence 10, Appl1	344
272	37.5	22.6	Sequence 36, Appl1	345
273	37	22.3	Sequence 16, Appl1	346
274	37	22.3	Sequence 10, Appl1	347
275	37	22.3	Sequence 10, Appl1	348
276	37	22.3	Sequence 2, Appl1	349
277	37	22.3	Sequence 6, Appl1	350
278	37	22.3	Sequence 8, Appl1	351
279	37	22.3	Sequence 34, Appl1	352
280	37	22.3	Sequence 35, Appl1	353
281	37	22.3	Sequence 37, Appl1	354
282	37	22.3	Sequence 37, Appl1	355
283	37	22.3	Sequence 37, Appl1	356
284	37	22.3	Sequence 37, Appl1	357
285	37	22.3	Sequence 37, Appl1	358
286	37	22.3	Sequence 37, Appl1	359
287	37	22.3	Sequence 37, Appl1	360
288	37	22.3	Sequence 37, Appl1	361
289	37	22.3	Sequence 37, Appl1	362
290	37	22.3	Sequence 37, Appl1	363
291	37	22.3	Sequence 37, Appl1	364
292	37	22.3	Sequence 37, Appl1	365
293	37	22.3	Sequence 37, Appl1	366
294	37	22.3	Sequence 37, Appl1	367
295	37	22.3	Sequence 37, Appl1	368
296	37	22.3	Sequence 37, Appl1	369
297	37	22.3	Sequence 37, Appl1	370
298	37	22.3	Sequence 37, Appl1	371
299	37	22.3	Sequence 37, Appl1	372
300	37	22.3	Sequence 37, Appl1	373
301	37	22.3	Sequence 37, Appl1	374
302	37	22.3	Sequence 37, Appl1	375
303	37	22.3	Sequence 37, Appl1	376
304	37	22.3	Sequence 37, Appl1	377
305	37	22.3	Sequence 37, Appl1	378
306	37	22.3	Sequence 37, Appl1	379
307	37	22.3	Sequence 37, Appl1	380
308	37	22.3	Sequence 37, Appl1	381
309	37	22.3	Sequence 37, Appl1	382
310	37	22.3	Sequence 37, Appl1	383
311	37	22.3	Sequence 37, Appl1	384
312	37	22.3	Sequence 37, Appl1	385
313	37	22.3	Sequence 37, Appl1	386
314	37	22.3	Sequence 37, Appl1	387
315	37	22.3	Sequence 37, Appl1	388
316	37	22.3	Sequence 37, Appl1	389
317	37	22.3	Sequence 37, Appl1	390
318	37	22.3	Sequence 37, Appl1	391
319	37	22.3	Sequence 37, Appl1	392

977 34 20.5 237 2 US-08-578-709-11 Sequence 11, Appl
978 34 20.5 240 3 US-08-908-332-9 Sequence 9, Appl
979 34 20.5 242 3 US-08-908-332-1 Sequence 1, Appl
980 34 20.5 247 4 US-09-370-838-112 Sequence 112, Appl
981 34 20.5 251 4 US-08-811-682-17 Sequence 17, Appl
982 34 20.5 261 4 US-09-134-001C-4193 Sequence 4193, Appl
983 34 20.5 262 4 US-09-008-271A-2 Sequence 2, Appl
984 34 20.5 269 4 US-09-724-864-47 Sequence 47, Appl
985 34 20.5 270 4 US-09-362-473-10 Sequence 10, Appl
986 34 20.5 284 2 US-09-055-095-1 Sequence 1, Appl
987 34 20.5 290 1 US-08-202-186-26 Sequence 26, Appl
988 34 20.5 310 2 US-08-701-191A-8 Sequence 8, Appl
989 34 20.5 312 2 US-09-047-026A-25 Sequence 25, Appl
990 34 20.5 315 3 US-09-135-639-4 Sequence 4, Appl
991 34 20.5 323 4 US-09-237-119-2 Sequence 2, Appl
992 34 20.5 343 2 US-08-599-171A-28 Sequence 28, Appl
993 34 20.5 343 2 US-08-646-590B-28 Sequence 28, Appl
994 34 20.5 343 3 US-09-069-226-28 Sequence 28, Appl
995 34 20.5 343 3 US-09-412-184-28 Sequence 28, Appl
996 34 20.5 347 1 US-08-461-244-3 Sequence 3, Appl
997 34 20.5 355 1 US-09-390-131-5 Sequence 5, Appl
998 34 20.5 360 1 US-08-450-393A-4 Sequence 4, Appl
999 34 20.5 360 3 US-09-043-627-2 Sequence 2, Appl
1000 34 20.5 360 4 US-08-446-669-4 Sequence 4, Appl

ALIGNMENTS

RESULT 1
US-09-045-284A-2
Sequence 2, Application US/09045284A
Patent No. 6255192
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107U51
CURRENT APPLICATION NUMBER: US/09/045,284A
CURRENT FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-045-284A-2

Query Match 100.0%; Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 SMRSGSFVQGLFGQHPDYFLMEPAHV 29
DB 50 SMRSGSFVQGLFGQHPDYFLMEPAHV 78

RESULT 2
US-09-190-911-1
Sequence 1, Application US/09190911
Patent No. 6365365
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerlich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CIP
CURRENT APPLICATION NUMBER: US/09/190,911
CURRENT FILING DATE: 1998-11-12
EARLIER APPLICATION NUMBER: 09/045,284
EARLIER FILING DATE: 1998-03-20

NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-09-190-911-1
Query Match 100.0%; Score 166; DB 4; Length 386;
Best Local Similarity 100.0%; Pred. No. 2e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SMRSGSFVQGLFGQHPDYFLMEPAHV 29
DB 50 SMRSGSFVQGLFGQHPDYFLMEPAHV 78

RESULT 3
US-09-263-023-2
Sequence 2, Application US/09263023
Patent No. 6037159
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kanagaki, Kenji
APPLICANT: Kanagaki, Reiichi
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263,023
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-263-023-2

Query Match 74.7%; Score 124; DB 3; Length 483;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;
QY 1 SMRSGSFVQGLFGQHPDYFLMEPAHV 29
DB 125 TWRSGSFVQGLFGQHPDYFLMEPAHV 153

RESULT 4
US-09-471-867-2
Sequence 2, Application US/09471867
Patent No. 6455289
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kanagaki, Kenji
APPLICANT: Kanagaki, Reiichi
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471,867
CURRENT FILING DATE: 1999-12-23
EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-09-471-867-2

Query Match 74.7%; Score 124; DB 4; Length 483;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFGQHPDYFLMEPMHV 29
Db 125 TWRSGSFFGELFNQNPVEFLYEPVHV 153

RESULT 5
US-09-263-023-4
Sequence 4, Application US/09263023
Patent No. 6037159
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Osami
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/263.023
CURRENT FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-263-023-4

Query Match 74.7%; Score 124; DB 3; Length 484;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFGQHPDYFLMEPMHV 29
Db 126 TWRSGSFFGELFNQNPVEFLYEPVHV 154

RESULT 6
US-09-471-867-4
Sequence 4, Application US/09471867
Patent No. 6455289
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kanagaki, Osami
APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/09/471.867
CURRENT FILING DATE: 1999-12-23

EARLIER APPLICATION NUMBER: US 09/263,023
EARLIER FILING DATE: 1999-03-05
EARLIER APPLICATION NUMBER: JP 10-54007
EARLIER FILING DATE: 1998-03-05
EARLIER APPLICATION NUMBER: JP 10-177844
EARLIER FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 4
LENGTH: 484
TYPE: PRT
ORGANISM: Homo sapiens
US-09-471-867-4

Query Match 74.7%; Score 124; DB 4; Length 484;
Best Local Similarity 69.0%; Pred. No. 6.2e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSGSFVGQLFGQHPDYFLMEPMHV 29
Db 126 TWRSGSFFGELFNQNPVEFLYEPVHV 154

RESULT 7
US-09-015-188-2
Sequence 2, Application US/09015188C
Patent No. 6393358
GENERAL INFORMATION:
APPLICANT: Williams, Kevin J
APPLICANT: Tabas, Ira
TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
FILE REFERENCE: JEFF-0231
CURRENT APPLICATION NUMBER: US/09/015,188C
CURRENT FILING DATE: 1998-01-29
NUMBER OF SEQ ID NOS: 17
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 411
TYPE: PRT
ORGANISM: Homo sapiens
US-09-015-188-2

Query Match 69.9%; Score 116; DB 4; Length 411;
Best Local Similarity 81.5%; Pred. No. 8.5e-10;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 RSGGSFVGQLFGQHPDYFLMEPMHV 29
Db 70 RSGGSFVGQLFGQHPDYFLMEPMHV 96

RESULT 8
US-08-655-878-2
Sequence 2, Application US/08655878
Patent No. 5827713
GENERAL INFORMATION:
APPLICANT: FUKUTA, MASAKAZU
APPLICANT: HABUCHI, OSAMI
TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESS:
STREET:
CITY:
STATE:
COUNTRY:
ZIP:
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3330 Hillview Avenue
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/900,491
FILING DATE: 25-JUL-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/443,865
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: LUTHER, BARBARA J.
REGISTRATION NUMBER: 33954
REFERENCE/DOCKET NUMBER: PF-0037 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-852-0195
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-900-491-6

Query Match 32.2%; Score 53.5; DB 1; Length 372;
Best Local Similarity 41.7%; Pred. No. 2.3;
Matches 10; Conservative 5; Mismatches 6; Indels 3; Gaps 1;

QY 1 SWRS---GSSFGOLFQHPDVEY 21
Db 308 SWDRRTGSIIFGLMMLPELFF 331

RESULT 12
US-09-202-491-7
Sequence 7, Application US/09202491
Patent No. 6432667
GENERAL INFORMATION:
APPLICANT: Valenzuela et al.
TITLE OF INVENTION: NOVEL LIGANDS, METHODS OF MAKING AND USES THEREOF
FILE REFERENCE: REG330-K
CURRENT APPLICATION NUMBER: US/09/202,491
CURRENT FILING DATE: 1998-11-16
EARLIER APPLICATION NUMBER: PCT/US97/10728
EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: 60/022,999
EARLIER FILING DATE: 1996-08-02
EARLIER APPLICATION NUMBER: 60/021,087
EARLIER FILING DATE: 1996-07-02
EARLIER APPLICATION NUMBER: 08/665,926
EARLIER FILING DATE: 1996-06-19
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 7
LENGTH: 496
TYPE: PRT
ORGANISM: Mus musculus
US-09-202-491-7

Query Match 28.3%; Score 47; DB 4; Length 496;
Best Local Similarity 43.5%; Pred. No. 31;
Matches 10; Conservative 2; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSFGOLFQHPDVEYLMPEAW 27

Db 356 GNEFVSQTLGQHRVYKIQKDW 378

RESULT 13
US-08-737-825-6
Sequence 6, Application US/08737825
Patent No. 5871922
GENERAL INFORMATION:
APPLICANT: SALMOND, GEORGE PEACOCK COPELAND
APPLICANT: MCGOWAN, SIMON JAMES
APPLICANT: SEBAIRIA, MOHAMMED
APPLICANT: COX, ANTHONY RICHARD JOHN
APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY
APPLICANT: PORTER, LAUREN ELIZABETH
APPLICANT: BYCROFT, BARRIE WALSHAM
APPLICANT: WILLIAMS, PAUL
APPLICANT: STEWART, GORDON SIDNEY ANDERSON BIRNIE
TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP
STREET: P.O. Box 747
CITY: Falls Church
STATE: Virginia
COUNTRY: USA
ZIP: 22040-0747
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: MICROSOFT WORD97
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/737,825
FILING DATE: 03-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SVENSSON, LEONARD R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 1009-0105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 376 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Erwinia carotovora
US-08-737-825-6

Query Match 28.0%; Score 46.5; DB 2; Length 376;
Best Local Similarity 36.7%; Pred. No. 27;
Matches 11; Conservative 8; Mismatches 10; Indels 1; Gaps 1;

QY 1 SWRSFGOLFQHPDVEY-LMPEAWV 29
Db 298 SLRAGRTWVGOLYGLDHTYLSLQGTGRV 327

RESULT 14
US-09-134-001C-5331
Sequence 5331, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS

FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 5331
LENGTH: 165
TYPE: PR1
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-5331

Query Match 27.1%; Score 45; DB 4; Length 165;
Best Local Similarity 61.5%; Pred. No. 18;
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 15 QHPDVEYLMPEAW 27
DB 59 QEPVFTAMEPSW 71

RESULT 15
US-08-551-356-4
Sequence 4, Application US/085513356
Patent No. 5830700
GENERAL INFORMATION:
APPLICANT: Iran1, Meher
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/551,356
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/349,762
FILING DATE:
APPLICATION NUMBER: US/07/998,271
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Holly, Julie A
REGISTRATION NUMBER: 33-246
REFERENCE/DOCKET NUMBER: 92-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-551-356-4

Query Match 27.1%; Score 45; DB 2; Length 643;
Best Local Similarity 52.9%; Pred. No. 83;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHP 17
DB 333 SWRSGSGTGTGNQNP 349

DB 333 SWRSGSGTGTGNQNP 349

RESULT 16
PCT-US93-12687-4
Sequence 4, Application PC/TUS9312687
GENERAL INFORMATION:
APPLICANT: Iran1, Meher H.
TITLE OF INVENTION: HYBRID CROSS-LINKING PROTEINS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/12687
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/998,271
FILING DATE: 31-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Parker, Gary E
REGISTRATION NUMBER: 31-648
REFERENCE/DOCKET NUMBER: 92-26PC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-547-8080 ext 322
TELEFAX: 206-548-2329
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 643 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12687-4

Query Match 27.1%; Score 45; DB 5; Length 643;
Best Local Similarity 55.2.9%; Pred. No. 83;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHP 17
DB 333 SWRSGSGTGTGNQNP 349

RESULT 17
US-08-206-176-2
Sequence 2, Application US/08206176
Patent No. 5639940
GENERAL INFORMATION:
APPLICANT: Garner, Ian
APPLICANT: Dalrymple, Michael A
APPLICANT: Prunkard, Donna E
APPLICANT: Foster, Donald C
TITLE OF INVENTION: Production of fibrinogen in transgenic
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zymogenetics, Inc.
STREET: 4225 Roosevelt Way, N.E.
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98105
COMPUTER READABLE FORM:

SEQUENCE CHARACTERISTICS:
LENGTH: 1336 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-12687-6

Query Match 27.1%; Score 45; DB 5; Length 1336;
Best Local Similarity 52.9%; Pred. No. 1.9e+02;
Matches 9; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGFQGHF 17
DB 1026 SWNSGSGGTGTCGNQNP 1042

RESULT 25
US-09-357-251-22
Sequence 22, Application US/09357251
Patent No. 6271441
GENERAL INFORMATION:
APPLICANT: Falco, S. Carl
APPLICANT: Farnou, Layo O.
APPLICANT: Orozco, Buddy
APPLICANT: Schwaber, James S.
TITLE OF INVENTION: Plant Aminoacyl-tRNA Synthetase
FILE REFERENCE: BB-1193
CURRENT APPLICATION NUMBER: US/09/357,251
EARLIER FILING DATE: 1999-07-20
EARLIER APPLICATION NUMBER: 60/093,530
NUMBER OF SEQ ID NOS: 37
SOFTWARE: Microsoft Office 97
SEQ ID NO: 22
LENGTH: 123
TYPE: PRT
ORGANISM: Glycine max
US-09-357-251-22

Query Match 26.5%; Score 44; DB 4; Length 123;
Best Local Similarity 40.0%; Pred. No. 19;
Matches 12; Conservative 2; Mismatches 8; Indels 8; Gaps 2;

QY 2 WRSGSFVQGFQGHF---DVFYLMPEA 26
DB 63 WNFDTLFOQ---QHAPDAHDTFLSEPA 89

RESULT 26
US-09-052-877-2
Sequence 2, Application US/09052877
Patent No. 6190912
GENERAL INFORMATION:
APPLICANT: Alnemri, Emad S.
TITLE OF INVENTION: BLK GENES, GENE PRODUCTS AND USES
TITLE OF INVENTION: THEREOF IN APOPTOSIS
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/052,877
FILING DATE: 31-MAR-1998
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 480140.433
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 662-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 150 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-052-877-2

Query Match 26.5%; Score 44; DB 4; Length 150;
Best Local Similarity 37.9%; Pred. No. 23;
Matches 11; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGFQGHFVFLYLMPEAHV 29
DB 119 AWSPDQDPQGFVFLYLMPEAHV 147

RESULT 27
US-08-740-223A-15
Sequence 15, Application US/08740223A
Patent No. 6265564
GENERAL INFORMATION:
APPLICANT: Davls, et al.
TITLE OF INVENTION: Expressed ligand - Vascular
TITLE OF INVENTION: Intercellular Signalling Molecule
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Regeneron Pharmaceuticals, Inc.
STREET: 777 Old Saw Mill Road
CITY: Tarrytown
STATE: NY
COUNTRY: USA
ZIP: 10591
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/740,223A
FILING DATE: 25-OCT-1996
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: USSN 60/022/959
FILING DATE: 02-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Coibert, Robert J.
REGISTRATION NUMBER: 36,108
REFERENCE/DOCKET NUMBER: REG 333
TELECOMMUNICATION INFORMATION:
TELEPHONE: 914-345-7721
TELEFAX: 914-345-7400
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 496 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: TIT2
LOCATION: 1...496
OTHER INFORMATION: mouse TIE-2 ligand 2
US-08-740-223A-15
Query Match 26.5%; Score 44; DB 4; Length 496;

Best Local Similarity 34.8%; Pred. No. 88;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSFFVGLFGQHPDVLYLMEPAW 27
Db 356 GNEFISQITGQHRVYLKIQKDW 378

RESULT 28

US-09-709-188-15

; Sequence 15, Application US/09709188
; Patent No. 6441137
; GENERAL INFORMATION:
; APPLICANT: Davis et al.
; TITLE OF INVENTION: Expressed Ligand - Vascular Intercellular Signaling Molecule
; FILE REFERENCE: REG 333-2
; CURRENT APPLICATION NUMBER: US/09/709,188
; CURRENT FILING DATE: 2000-11-09
; PRIOR APPLICATION NUMBER: 08/740,223
; PRIOR FILING DATE: 1996-10-25
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 15
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Mus sp.
US-09-709-188-15

Query Match 26.5%; Score 44; DB 4; Length 496;
Best Local Similarity 34.8%; Pred. No. 88;
Matches 8; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 5 GSSFFVGLFGQHPDVLYLMEPAW 27
Db 356 GNEFISQITGQHRVYLKIQKDW 378

RESULT 29

US-08-295-814E-4

; Sequence 4, Application US/08295814E
; Patent No. 5658786
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E.
; APPLICANT: Borden, Laurence A.
; APPLICANT: Hartig, Paul R.
; APPLICANT: Weinstank, Richard L.
; TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
; TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/295,814E
FILING DATE: DECEMBER 19, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 4058-B-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 4;

SEQUENCE CHARACTERISTICS:

LENGTH: 627 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-295-814E-4

Query Match 26.5%; Score 44; DB 1; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 2 WR-----GSSFFVGLFGQHPDVLYLMEPAW 29

Db 117 WRVCPLEFGIGYATQVIEAHNLVYIITLAWAI 150

RESULT 30

US-08-291-299-10

; Sequence 10, Application US/08291299
; Patent No. 5766848
; GENERAL INFORMATION:
; APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINE/GABA
; TITLE OF INVENTION: TRANSPORTER AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,299
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 42897-A/JPM/TEP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 977-9550
TELEFAX: (212) 664-0525
TELEX: 422523 COOP UT
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 627 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-291-299-10

Query Match 26.5%; Score 44; DB 1; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

QY 2 WR-----GSSFFVGLFGQHPDVLYLMEPAW 29

Db 117 WRVCPLEFGIGYATQVIEAHNLVYIITLAWAI 150

RESULT 31

US-09-343-361-4

; Sequence 4, Application US/09343361
; Patent No. 6225115
; GENERAL INFORMATION:
; APPLICANT: Smith, Kelli E. et al
; TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses


```

? TITLE OF INVENTION: Theoeof
? FILE REFERENCE: 40558-D
? CURRENT APPLICATION NUMBER: US-09/343,361
? CURRENT FILING DATE: 1999-06-30
? NUMBER OF SEQ ID NOS: 26
? SOFTWARE: PatentIn Ver. 2.0 - beta
? SEQ ID NO. 4: beta
? LENGTH: 627
? TYPE: PRT
? ORGANISM: Rattus norvegicus
? OS-09-343-361-4

```

	Query Match	Best Local Similarity	Score	DB 4	Length
Matches	9	Conservative	7	Mismatches	12
Indels					6
Gaps					

```

RESULT 32
PCT-US93-01959-4
: Sequence 4, Application PC/TUS9301959
: GENERAL INFORMATION:
: APPLICANT: Smith, E. Kelli
: APPLICANT: Borden, A. Laurence
: APPLICANT: Hartig, R. Paul
: APPLICANT: Weltschank, L. Richard
: TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
: TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Cooper & Dunham
: STREET: 30 Rockefeller Plaza
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10112
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Releasee #1.24
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US93/01959
: FILING DATE: 19930304
: CLASSIFICATION:
: ATTORNEY/AGENT INFORMATION:
: NAME: White, John
: REGISTRATION NUMBER: 28,678
: REFERENCE/DOCKET NUMBER: 40558A
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-977-9550
: TELEFAX: 212-664-0525
: TELEX: 422523 COOP UI
: INFORMATION FOR SEQ ID NO: 4:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 627 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US93-01959-4

```

```

Query Match: 26.5%; Score 44; DB 5; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,le+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1
QY 2 WR-----SSSPVQGLFSCQHPDVEYFLMEPAHV 29
      |||  |||  |||  |||  |||  |||  |||  |||
Db 117 WRRVCPDLEIGATGVIEAHLNYYVITLLAMAV 150

```

```

1  RESULT 33
2  PCT-US95-10579-10
3  Sequence 10 Application PC/TUS9510579
4  GENERAL INFORMATION:
5  APPLICANT: SYNAPTIC PHARMACEUTICAL CORPORATION
6  TITLE OF INVENTION: DNA ENCODING A HUMAN BETAINE/GABA
7  TITLE OF INVENTION: TRANSPORTER AND USES THEREOF
8  NUMBER OF SEQUENCES: 10
9  CORRESPONDENCE ADDRESS:
10 ADDRESS: Cooper & Dunham LLP
11 STREET: 1185 Avenue of the Americas
12 CITY: New York
13 STATE: New York
14 COUNTRY: U.S.A.
15 ZIP: 10036
16
17 COMPUTER READABLE FORM:
18 MEDIUM TYPE: Floppy disk
19 COMPUTER: IBM PC compatible
20 OPERATING SYSTEM: PC-DOS/MS-DOS
21 SOFTWARE: PatentIn Release #1.24
22 CURRENT APPLICATION DATA:
23 APPLICATION NUMBER: PCT/US95/10579
24 FILING DATE:
25
26 CLASSIFICATION:
27 ATTORNEY/AGENT INFORMATION:
28 NAME: White, John P.
29 REGISTRATION NUMBER: 28,678
30 REFERENCE/DOCKET NUMBER: 42897-A-PCT/JPW/MAT
31 TELECOMMUNICATION INFORMATION:
32 TELEPHONE: (212) 278-0400
33 TELEFAX: (212) 391-0525
34 INFORMATION FOR SEQ ID NO: 10:
35 SEQUENCE CHARACTERISTICS:
36 LENGTH: 627 amino acids
37 TYPE: amino acid
38 TOPOLOGY: linear
39 MOLECULE TYPE: protein
40
41 PCT-US95-10579-10

```

```

Query Match      26.5%; Score 44; DB 5; Length 627;
Best Local Similarity 26.5%; Pred. No. 1,1e+02;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1

OY      2 WR-----SSSSVGLGCGHPDVTYLMPPAHV 29
          ||      |      |      |      |
Db      117 WRKVCPLFGESIGATATVTEAHNLNVITIIILAAAI 150

```

RESULT 34
 US-08-295-814E-10
 Sequence 10. Application US/06295814E
 Patent No. 5658786
 GENERAL INFORMATION:
 APPLICANT: Smith, Kelli E.
 APPLICANT: Borden, Laurence A.
 APPLICANT: Hartley, Paul R.
 APPLICANT: Weinstein, Richard L.
 TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
 TITLE OF INVENTION: TRANSPORTERS AND USES THEREOF
 NUMBER OF SPOUNCES: 35
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.30
 CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/295,814E
FILING DATE: DECEMBER 19, 1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40558-B-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 632 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-295-814E-10

Query Match
Best Local Similarity 26.5%; Score 44; DB 1; Length 632;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

OY 2 WR-----SSSFVGGQLFGQHPDVEYLMPEAMHV 29
DB 122 WRKVCPLFEGIGYATQVTEAHNLNVIITLLAMAI 155

RESULT 35
US-09-343-361-10
Sequence 10, Application US/09343361
Patent No. 6225115
GENERAL INFORMATION:
APPLICANT: Smith, Kelli E. et al
TITLE OF INVENTION: DNA Encoding Taurine and GABA Transporters and Uses
FILE REFERENCE: 40558-D
CURRENT APPLICATION NUMBER: US/09/343,361
FILING DATE: 1999-06-30
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn Ver. 2.0 - beta
SEQ ID NO 10
LENGTH: 632
TYPE: PRT
ORGANISM: Homo sapiens
US-09-343-361-10

Query Match
Best Local Similarity 26.5%; Score 44; DB 4; Length 632;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

OY 2 WR-----SSSFVGGQLFGQHPDVEYLMPEAMHV 29
DB 122 WRKVCPLFEGIGYATQVTEAHNLNVIITLLAMAI 155

PCT-US93-01959-10
Sequence 10, Application PC/TUS9301959
GENERAL INFORMATION:
APPLICANT: Smith, E. Kelli
APPLICANT: Borden, A. Laurence
APPLICANT: Hartig, R. Paul
APPLICANT: Weinschenk, L. Richard
TITLE OF INVENTION: DNA ENCODING TAURINE AND GABA
TRANSPORTERS AND USES THEREOF
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSER: Cooper & Dunham
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10112

```

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.24
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/01959
FILING DATE: 19930304
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: White, John
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 40558A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-977-9550
TELEFAX: 422523 COOP UI
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 632 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US93-01959-10

Query Match
Best Local Similarity 26.5%; Score 44; DB 5; Length 632;
Matches 9; Conservative 7; Mismatches 12; Indels 6; Gaps 1;

OY 2 WR-----SSSFVGGQLFGQHPDVEYLMPEAMHV 29
DB 122 WRKVCPLFEGIGYATQVTEAHNLNVIITLLAMAI 155

RESULT 37
US-08-095-734-2
Sequence 2, Application US/08095734
Patent No. 5807723
GENERAL INFORMATION:
APPLICANT: Aldovini, Anna
APPLICANT: Young, Richard A.
TITLE OF INVENTION: Homologously Recombinant Slow Growing
MYCOBACTERIA AND USES THEREFOR
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSER: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/095,734
FILING DATE: 22-JUL-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI93-11M
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1271 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

```

MOLECULE TYPE: protein
US-08-095-734-2

Query Match 26.5%; Score 44; DB 1; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;
DY 2 WRSQSFVQLFGQHPDVFYIMEPA 26
DB 116 WRTGSPTPGARGHPPEAIDLAEDA 140

RESULT 38

US-08-444-623-2
Sequence 2, Application US/08444623
Patent No. 586403
GENERAL INFORMATION:
APPLICANT: Aldovini, Anna
APPLICANT: Young, Richard A.
TITLE OF INVENTION: Homologously Recombinant Slow Growing
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/444,623
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/095,734
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: 07/711,334
FILING DATE: 06-JUN-1991
APPLICATION NUMBER: 07/367,894
FILING DATE: 19-JUN-1989
APPLICATION NUMBER: PCT/US90/03451
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: PCT/US89/02962
FILING DATE: 07-JUL-1989
APPLICATION NUMBER: 07/361,944
FILING DATE: 05-JUN-1989
APPLICATION NUMBER: 07/223,089
FILING DATE: 22-JUL-1988
APPLICATION NUMBER: 07/216,390
FILING DATE: 07-JUL-1988
APPLICATION NUMBER: 07/163,546
FILING DATE: 03-MAR-1988
APPLICATION NUMBER: PCT/US88/00614
FILING DATE: 29-FEB-1988
APPLICATION NUMBER: 07/020,451
FILING DATE: 02-MAR-1987
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH193-11M2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-861-6240
TELEFAX: 617-861-9540
INFORMATION FOR SEQ. ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1271 amino acids
TYPE: amino acid
STRANDEDNESS: single

TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-444-623-2

Query Match 26.5%; Score 44; DB 2; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;
DY 2 WRSQSFVQLFGQHPDVFYIMEPA 26
DB 116 WRTGSPTPGARGHPPEAIDLAEDA 140

RESULT 39

US-08-471-869-2
Sequence 2, Application US/08471869
Patent No. 6022745
GENERAL INFORMATION:
APPLICANT: Aldovini, Anna
APPLICANT: Young, Richard A.
TITLE OF INVENTION: Homologously Recombinant Slow Growing
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Milltia Drive
CITY: Lexington
STATE: MA
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,869
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/08267
FILING DATE: 22-JUL-1994
APPLICATION NUMBER: 08/095,734
FILING DATE: 22-JUL-1993
APPLICATION NUMBER: 07/711,334
FILING DATE: 06-JUN-1991
APPLICATION NUMBER: 07/367,894
FILING DATE: 19-JUN-1989
APPLICATION NUMBER: PCT/US90/03451
FILING DATE: 18-JUN-1990
APPLICATION NUMBER: PCT/US89/02962
FILING DATE: 07-JUL-1989
APPLICATION NUMBER: 07/361,944
FILING DATE: 05-JUN-1989
APPLICATION NUMBER: 07/223,089
FILING DATE: 22-JUL-1988
APPLICATION NUMBER: 07/216,390
FILING DATE: 07-JUL-1988
APPLICATION NUMBER: 07/163,546
FILING DATE: 03-MAR-1988
APPLICATION NUMBER: PCT/US88/00614
FILING DATE: 29-FEB-1988
PRIOR APPLICATION DATA:

```

; APPLICATION NUMBER: US 07/020,451
; FILING DATE: 02-MAR-1987
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; REFERENCE/DOCKET NUMBER: WH193-11MA2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-471-869-2

Query Match      26.5%; Score 44; DB 3; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY      2 WRGSSFFVGLFGQHPDVLYLMEPA 26
DB      116 WRGSPTPGPARGHPPEAIDLAEDA 140

RESULT 40
; US-09-342-563-2
; Sequence 2, Application US/09342563
; Patent No. 6355486
; GENERAL INFORMATION:
; APPLICANT: Aldovini, Anna
; APPLICANT: Young, Richard A.
; TITLE OF INVENTION: Homologously Recombinant Slow Growing
; TITLE OF INVENTION: Mycobacteria and Uses Therefor
; FILE REFERENCE: WH193-11MA3
; CURRENT APPLICATION NUMBER: US/09/342,563
; EARLIER FILING DATE: 1999-06-29
; EARLIER APPLICATION NUMBER: US 08/471,869
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: US 08/095,734
; EARLIER FILING DATE: 1993-07-22
; EARLIER APPLICATION NUMBER: PCT/US94/08267
; EARLIER FILING DATE: 1994-07-22
; EARLIER APPLICATION NUMBER: US 07/11,334
; EARLIER FILING DATE: 1991-06-06
; EARLIER APPLICATION NUMBER: US 07/367,894
; EARLIER FILING DATE: 1989-06-19
; EARLIER APPLICATION NUMBER: PCT/US90/03451
; EARLIER FILING DATE: 1990-06-18
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 2
; LENGTH: 1271
; TYPE: PRT
; ORGANISM: Mycobacterium bovis Bacille Calmette-Guerin
; US-09-342-563-2

Query Match      26.5%; Score 44; DB 4; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY      2 WRGSSFFVGLFGQHPDVLYLMEPA 26
DB      116 WRGSPTPGPARGHPPEAIDLAEDA 140

RESULT 41
; PCT-US94-08267-2
; Sequence 2, Application PC/TUS9408267
; GENERAL INFORMATION:
; APPLICANT:
```

```

; TITLE OF INVENTION: Homologously Recombinant Slow Growing
; TITLE OF INVENTION: Mycobacteria and Uses Therefor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Millitia Drive
; CITY: Lexington
; STATE: MA
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/08267
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/095,734
; FILING DATE: 22-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Granahan, Patricia
; REGISTRATION NUMBER: 32,227
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1271 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US94-08267-2

Query Match      26.5%; Score 44; DB 5; Length 1271;
Best Local Similarity 40.0%; Pred. No. 2.5e+02;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY      2 WRGSSFFVGLFGQHPDVLYLMEPA 26
DB      116 WRGSPTPGPARGHPPEAIDLAEDA 140

RESULT 42
; US-08-469-005A-10
; Sequence 10, Application US/08469005A
; Patent No. 5665874
; GENERAL INFORMATION:
; APPLICANT: KUHADA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
```

APPLICATION NUMBER: 08/188,426
FILING DATE: 24-JAN-1994
APPLICATION NUMBER: 08/096,908
FILING DATE: 26-JUL-1993
APPLICATION NUMBER: 07/917,716
FILING DATE: 24-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: Posorske, Laurence H
REGISTRATION NUMBER: 34,638
REFERENCE/DOCKET NUMBER: 062482-0113
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-639-7700
TELEFAX: 202-639-7890
TELEX:
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 2509 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
US-08-469-005A-10

Query Match 26.5%; Score 44; DB 1; Length 2509;
Best Local Similarity 26.3%; Pred. No. 5.4e+02;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SFFVGLFGQHPDVFYIMEP 25
DB 40 WKAGLYGLPRSGKLKLDLSRFASFFGVHPKQAHMDP 77

RESULT 43
US-09-261-907-2
Sequence 2, Application US/09261907A
Patent No. 6294364
GENERAL INFORMATION:
APPLICANT: ELLIS, CATHERINE
APPLICANT: LONSDALE, JOHN
APPLICANT: BERGSMAN, DEBK J.
APPLICANT: MOONEY, JEFFREY L.
APPLICANT: DEPIERA, MEGAN E.
APPLICANT: CHAPMAN, CONRAD
TITLE OF INVENTION: HUMAN PAS
FILE REFERENCE: GP-70603
CURRENT APPLICATION NUMBER: US/09/261.907A
CURRENT FILING DATE: 1999-03-03
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2511
TYPE: PRT
ORGANISM: HOMO SAPIENS
US-09-261-907-2

Query Match 26.5%; Score 44; DB 4; Length 2511;
Best Local Similarity 26.3%; Pred. No. 5.4e+02;
Matches 10; Conservative 4; Mismatches 10; Indels 14; Gaps 1;

QY 2 WRSG-----SFFVGLFGQHPDVFYIMEP 25
DB 40 WKAGLYGLPRSGKLKLDLSRFASFFGVHPKQAHMDP 77

RESULT 44
US-08-906-769-89
Sequence 89, Application US/08906769
Patent No. 6077687
GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,769
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-769-89

Query Match 25.9%; Score 43; DB 3; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWRGSSFFVGLFGQHPDVFYIMEPAW 27
DB 75 SVRVGTSPGRRGVSHPKQAHMDP 101

RESULT 45
US-08-906-616-89
Sequence 89, Application US/08906616
Patent No. 6121035
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA AMINOPEPTIDASE PROTEINS AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA

ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,616
FILING DATE: 05-AUG-1997
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2-3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-616-89

Query Match 25.9%; Score 43; DB 3; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SRRGSSFTVGLFGQHPDVFYLMPEAW 27
DB 75 SVRVGTSFGGRGSRVHPXAQIIKHPAY 101

RESULT 46
US-08-817-795-89
Sequence 89, Application US/08817795
Patent No. 6139840
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Hunter, Shirley Wu
APPLICANT: Frank, Glenn R.
APPLICANT: Heath, Andrew W.
APPLICANT: Yamaka, Miles Yamanaka
APPLICANT: Arfsten, Ann
APPLICANT: Dale, Beverly
APPLICANT: Stiegler, Gary
TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES, AND USES THEREOF
NUMBER OF SEQUENCES: 119
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/817,795
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/14442
FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: Gary J. Connell
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-817-795-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SRRGSSFTVGLFGQHPDVFYLMPEAW 27
DB 75 SVRVGTSFGGRGSRVHPXAQIIKHPAY 101

RESULT 47
US-08-639-075A-89
Sequence 89, Application US/08639075A
Patent No. 6150125
GENERAL INFORMATION:
APPLICANT: Grieve, Robert B.
APPLICANT: Rushlow, Keith E.
APPLICANT: Wu Hunter, Shirley
APPLICANT: Frank, Glenn R.
APPLICANT: Stiegler, Gary
APPLICANT: Gaines, Patrick J.
APPLICANT: Silver, Gary
TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
TITLE OF INVENTION: MOLECULES AND USES THEREOF
NUMBER OF SEQUENCES: 190
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/639,075A
FILING DATE: 24-APR-1996
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-639-075A-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFYLMPEAW 27
Db 75 SVRVGTSGFQGRGVSHPXAQITKHPAY 101

RESULT 48

US-09-012-431-89
Sequence 89, Application US/09012431
Patent No. 6180383

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.
Rushlow, Keith E.
Wu Hunter, Shirley
Frank, Glenn R.
Stiegler, Gary
Gaines, Patrick J.
Silver, Gary

TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID
MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,431

FILING DATE: 23-Jan-1998

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/639,075

FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 256 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

SEQUENCE DESCRIPTION: SEQ ID NO: 89:

US-09-012-431-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFYLMPEAW 27
Db 75 SVRVGTSGFQGRGVSHPXAQITKHPAY 101

RESULT 49

US-09-012-692-89
Sequence 89, Application US/09012692
Patent No. 6214579

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.

APPLICANT: Rushlow, Keith E.

APPLICANT: Wu Hunter, Shirley

APPLICANT: Frank, Glenn R.

APPLICANT: Stiegler, Gary
Gaines, Patrick J.

APPLICANT: Silver, Gary

TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID

TITLE OF INVENTION: MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/012,692

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/639,075

FILING DATE: 24-APR-1996

ATTORNEY/AGENT INFORMATION:

NAME: Connell, Gary J.

REGISTRATION NUMBER: 32,020

REFERENCE/DOCKET NUMBER: 2618-25-C2

TELECOMMUNICATION INFORMATION:

TELEPHONE: (303) 863-9700

TELEFAX: (303) 863-0223

INFORMATION FOR SEQ ID NO: 89:

SEQUENCE CHARACTERISTICS:

LENGTH: 256 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-09-012-692-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFYLMPEAW 27
Db 75 SVRVGTSGFQGRGVSHPXAQITKHPAY 101

RESULT 50

US-08-906-613-89
Sequence 89, Application US/08906613
Patent No. 6232096

GENERAL INFORMATION:

APPLICANT: Grieve, Robert B.

APPLICANT: Rushlow, Keith E.

APPLICANT: Wu Hunter, Shirley

APPLICANT: Frank, Glenn R.

APPLICANT: Stiegler, Gary

APPLICANT: Gaines, Patrick J.

APPLICANT: Silver, Gary

TITLE OF INVENTION: FLEA PROTEASE PROTEINS, NUCLEIC ACID

TITLE OF INVENTION: MOLECULES AND USES THEREOF

NUMBER OF SEQUENCES: 190

CORRESPONDENCE ADDRESS:

ADDRESSEE: Sheridan Ross & McIntosh

STREET: 1700 Lincoln Street, Suite 3500

CITY: Denver

STATE: Colorado

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/906,613
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/639,075
FILING DATE: 24-APR-1996
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 2618-25-C2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (303) 863-9700
TELEFAX: (303) 863-0223
INFORMATION FOR SEQ ID NO: 89:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-906-613-89

Query Match 25.9%; Score 43; DB 4; Length 256;
Best Local Similarity 37.0%; Pred. No. 60;
Matches 10; Conservative 4; Mismatches 13; Indels 0; Gaps 0;
QY 1 SWRGSFVQLEFGHPDVEYLMPAW 27
DB 75 SVRVGTSGRGSGVHPXAGITKHPAY 101

Search completed: February 20, 2003, 13:35:00
Job time : 49.1429 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:33:52 ; Search time 13.2571 Seconds
(without alignments)
55.888 Million cell updates/sec

Title: US-09-816-825-2_COPY_50_78

Perfect score: 166
Sequence: 1 SWRSGSEVGFQGHDPVFLMEPMHV 29

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 25548876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :
Published Applications: AA: *
1: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep: *
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep: *
3: /cgn2_6/ptodata/1/pubppaa/US06_NEW_PUB.pep: *
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep: *
5: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep: *
6: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep: *
7: /cgn2_6/ptodata/1/pubppaa/PCYUS_PUBCOMB.pep: *
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep: *
9: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep: *
10: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep: *
11: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep: *
12: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep: *
13: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep: *
14: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	166	100.0	171	10	US-09-827-602-8	Sequence 8, Appl1
2	166	100.0	386	9	US-10-007-262-1	Sequence 1, Appl1
3	166	100.0	386	10	US-09-816-825-2	Sequence 2, Appl1
4	160	96.4	169	10	US-09-927-602-6	Sequence 6, Appl1
5	160	96.4	395	10	US-09-927-602-2	Sequence 2, Appl1
6	160	96.4	395	10	US-09-927-602-3	Sequence 3, Appl1
7	160	96.4	418	10	US-09-927-602-5	Sequence 5, Appl1
8	157	94.6	169	10	US-09-927-602-7	Sequence 7, Appl1
9	157	94.6	390	10	US-09-927-602-4	Sequence 4, Appl1
10	124	74.7	169	10	US-09-927-602-9	Sequence 9, Appl1
11	124	74.7	483	9	US-10-212-933-2	Sequence 2, Appl1
12	124	74.7	484	9	US-09-833-790-255	Sequence 255, Appl1
13	124	74.7	531	10	US-09-927-602-10	Sequence 10, Appl1
14	116	69.9	174	10	US-09-927-602-11	Sequence 11, Appl1
15	101	60.8	174	9	US-09-738-626-5973	Sequence 5973, Appl1
16	50	30.1	344	10	US-09-888-623-2	Sequence 2, Appl1
17	49	29.5	265	10	US-09-888-623-14	Sequence 14, Appl1
18	49	29.5	265	10	US-09-888-623-16	Sequence 16, Appl1
19	49	29.5	318	10	US-09-888-623-16	Sequence 16, Appl1

20	48	28.9	460	10	US-09-815-242-5057	Sequence 5057, App
21	48	28.9	753	10	US-09-888-615-68	Sequence 68, Appl1
22	47	28.3	701	10	US-09-815-242-12050	Sequence 12050, A
23	47	28.3	925	9	US-09-738-626-3608	Sequence 3608, App
24	46	27.7	302	9	US-09-738-626-3746	Sequence 3746, App
25	45.5	27.4	844	10	US-09-272-809-6	Sequence 6, Appl1
26	45	27.1	99	10	US-09-864-761-41772	Sequence 41772, A
27	45	27.1	239	10	US-09-925-299-834	Sequence 834, App
28	45	27.1	376	9	US-09-738-626-6189	Sequence 6189, App
29	45	27.1	791	10	US-09-735-705-170	Sequence 170, App
30	45	27.1	791	10	US-09-850-716A-170	Sequence 170, App
31	45	27.1	791	10	US-09-897-778-170	Sequence 170, App
32	45	27.1	847	9	US-10-112-527-4	Sequence 4, Appl1
33	45	27.1	881	10	US-09-850-716A-430	Sequence 430, App
34	45	27.1	881	10	US-09-897-778-430	Sequence 430, App
35	45	27.1	920	10	US-09-735-705-357	Sequence 357, App
36	45	27.1	920	10	US-09-850-716A-357	Sequence 357, App
37	45	27.1	920	10	US-09-897-778-357	Sequence 357, App
38	45	27.1	942	9	US-09-974-298-71	Sequence 71, Appl1
39	45	27.1	942	10	US-09-919-172-87	Sequence 87, Appl1
40	45	27.1	943	10	US-09-735-705-161	Sequence 161, App
41	45	27.1	943	10	US-09-850-716A-161	Sequence 161, App
42	45	27.1	943	10	US-09-897-778-161	Sequence 161, App
43	44	26.5	812	10	US-09-815-242-11294	Sequence 11294, A
44	43.5	25.2	1054	9	US-09-738-626-5852	Sequence 5852, App
45	43	25.9	466	10	US-09-925-302-729	Sequence 729, App
46	43	25.9	1160	10	US-09-815-242-5480	Sequence 5480, App
47	43	25.9	1168	10	US-09-815-242-12151	Sequence 12151, A
48	42.5	25.6	80	9	US-09-764-870-510	Sequence 510, App
49	42.5	25.6	94	10	US-10-001-835-184	Sequence 184, App
50	42.5	25.6	229	9	US-09-992-598-281	Sequence 281, App
51	42.5	25.6	229	9	US-09-989-723-281	Sequence 281, App
52	42.5	25.6	229	9	US-09-989-735-281	Sequence 281, App
53	42.5	25.6	229	9	US-09-989-735-281	Sequence 281, App
54	42.5	25.6	229	9	US-09-989-730-281	Sequence 281, App
55	42.5	25.6	229	9	US-09-990-436-281	Sequence 281, App
56	42.5	25.6	229	9	US-09-991-181-281	Sequence 281, App
57	42.5	25.6	229	9	US-09-993-687-281	Sequence 281, App
58	42.5	25.6	229	9	US-09-989-734-281	Sequence 281, App
59	42.5	25.6	229	9	US-09-997-653-281	Sequence 281, App
60	42.5	25.6	229	9	US-09-993-667-281	Sequence 281, App
61	42.5	25.6	229	9	US-09-990-438-281	Sequence 281, App
62	42.5	25.6	229	9	US-09-990-562-281	Sequence 281, App
63	42.5	25.6	229	9	US-09-997-428-281	Sequence 281, App
64	42.5	25.6	229	9	US-09-997-666-281	Sequence 281, App
65	42.5	25.6	229	10	US-09-989-723-281	Sequence 281, App
66	42.5	25.6	229	10	US-09-989-723-281	Sequence 281, App
67	42.5	25.6	229	10	US-09-989-727-281	Sequence 281, App
68	42.5	25.6	229	10	US-09-989-727-281	Sequence 281, App
69	42.5	25.6	229	10	US-09-989-731-281	Sequence 281, App
70	42.5	25.6	229	10	US-09-989-732-281	Sequence 281, App
71	42.5	25.6	229	10	US-09-991-073-281	Sequence 281, App
72	42.5	25.6	229	10	US-09-990-442-281	Sequence 281, App
73	42.5	25.6	229	10	US-09-991-163-281	Sequence 281, App
74	42.5	25.6	229	10	US-09-993-604-281	Sequence 281, App
75	42.5	25.6	229	10	US-09-990-456-281	Sequence 281, App
76	42.5	25.6	229	10	US-09-989-721-281	Sequence 281, App
77	42	25.3	66	10	US-09-867-550-1170	Sequence 1170, App
78	42	25.3	74	9	US-09-989-920-255	Sequence 255, App
79	42	25.3	252	10	US-09-347-064-2	Sequence 2, Appl1
80	42	25.3	252	10	US-09-347-064-8	Sequence 8, Appl1
81	42	25.3	335	9	US-10-127-884-170	Sequence 170, App
82	42	25.3	334	9	US-09-924-400-304	Sequence 304, App
83	42	25.3	334	9	US-09-924-400-333	Sequence 333, App
84	42	25.3	384	9	US-09-924-400-340	Sequence 340, App
85	42	25.3	384	10	US-09-825-301-8	Sequence 8, Appl1
86	42	25.3	384	10	US-09-810-936-334	Sequence 334, App
87	42	25.3	384	10	US-09-810-936-333	Sequence 333, App
88	42	25.3	384	10	US-09-429-755-336	Sequence 336, App
89	42	25.3	334	9	US-09-924-400-334	Sequence 334, App
90	42	25.3	529	10	US-09-810-936-334	Sequence 334, App
91	42	25.3	529	10	US-09-924-400-305	Sequence 305, App
92	42	25.3	666	9	US-09-924-400-305	Sequence 305, App

93	42	25.3	656	9	US-10-012-896-379	Sequence 379, App	166	40	24.1	1491	10	US-09-815-242-5568	Sequence 5568, App
94	42	25.3	656	9	US-09-895-793-379	Sequence 379, App	167	40	24.1	1502	10	US-09-815-242-12162	Sequence 12162, A
95	42	25.3	656	9	US-09-895-814-379	Sequence 379, App	168	39.5	23.8	1248	10	US-09-825-992-958	Sequence 958, App
96	42	25.3	656	10	US-09-825-301-9	Sequence 9, Appli	169	39.5	23.8	323	10	US-09-903-395-2	Sequence 2, Appli
97	42	25.3	656	10	US-09-753-143-379	Sequence 379, App	170	38.5	23.8	333	10	US-09-828-113-33	Sequence 33, Appli
98	42	25.3	656	10	US-09-780-669-379	Sequence 379, App	171	39.5	23.8	337	9	US-09-860-846-14	Sequence 14, Appli
99	42	25.3	656	10	US-09-810-936-305	Sequence 305, App	172	39.5	23.8	337	10	US-09-861-889-14	Sequence 14, Appli
100	42	25.3	656	10	US-09-822-827-379	Sequence 379, App	173	39.5	23.8	500	9	US-09-712-363-774	Sequence 274, App
101	42	25.3	656	10	US-09-429-755-305	Sequence 305, App	174	39.5	23.8	877	9	US-10-043-418-1	Sequence 1, Appli
102	42	25.3	671	9	US-09-924-400-306	Sequence 306, App	175	39.5	23.8	1169	10	US-09-815-242-13448	Sequence 13448, A
103	42	25.3	671	9	US-10-012-896-380	Sequence 380, App	176	39.5	23.8	1166	10	US-09-815-242-13693	Sequence 13693, A
104	42	25.3	671	9	US-09-895-793-380	Sequence 380, App	177	39.5	23.8	3782	9	US-09-860-846-4	Sequence 4, Appli
105	42	25.3	671	9	US-09-895-814-380	Sequence 380, App	178	39.5	23.8	3782	10	US-09-861-889-4	Sequence 4, Appli
106	42	25.3	671	10	US-09-825-301-10	Sequence 10, Appli	179	39	23.5	17	9	US-09-996-634-91	Sequence 9, Appli
107	42	25.3	671	10	US-09-759-143-380	Sequence 380, App	180	39	23.5	63	10	US-09-864-761-42380	Sequence 42380, A
108	42	25.3	671	10	US-09-780-669-380	Sequence 380, App	181	39	23.5	111	10	US-09-925-300-980	Sequence 980, App
109	42	25.3	671	10	US-09-810-936-306	Sequence 306, App	182	39	23.5	154	10	US-09-764-864-1156	Sequence 1156, Ap
110	42	25.3	671	10	US-09-822-827-380	Sequence 380, App	183	39	23.5	154	10	US-09-764-864-1569	Sequence 1569, Ap
111	42	25.3	671	10	US-09-429-755-306	Sequence 306, App	184	39	23.5	155	10	US-09-729-674-104	Sequence 104, App
112	42	25.3	876	10	US-09-815-242-12623	Sequence 12623, A	185	39	23.5	171	9	US-09-895-913A-26	Sequence 26, App
113	42	25.3	1719	9	US-10-012-896-378	Sequence 378, App	186	39	23.5	178	10	US-09-864-761-41969	Sequence 41969, A
114	42	25.3	1719	9	US-09-895-793-378	Sequence 378, App	187	39	23.5	284	10	US-09-771-161A-106	Sequence 106, App
115	42	25.3	1719	9	US-09-895-814-378	Sequence 378, App	188	39	23.5	299	10	US-09-886-055-41	Sequence 41, Appli
116	42	25.3	1719	10	US-09-759-143-378	Sequence 378, App	189	39	23.5	300	10	US-09-919-770-2	Sequence 2, Appli
117	42	25.3	1719	10	US-09-780-669-378	Sequence 378, App	190	39	23.5	305	10	US-09-925-300-51205	Sequence 1205, Ap
118	42	25.3	1719	10	US-09-822-827-378	Sequence 378, App	191	39	23.5	314	10	US-09-206-576-2	Sequence 2, Appli
119	41.5	25.0	56	9	US-09-796-692-748	Sequence 748, App	192	39	23.5	318	10	US-09-350-368-16	Sequence 16, Appli
120	41.5	25.0	105	10	US-09-864-761-33632	Sequence 33632, A	193	39	23.5	322	10	US-09-942-447-5	Sequence 5, Appli
121	41.5	25.0	210	12	US-10-029-079-3	Sequence 3, Appli	194	39	23.5	365	10	US-09-818-512-4	Sequence 4, Appli
122	41.5	25.0	730	9	US-09-978-295A-102	Sequence 102, App	195	39	23.5	370	9	US-10-114-693-174	Sequence 2, Appli
123	41.5	25.0	730	9	US-09-978-697-102	Sequence 102, App	196	39	23.5	370	10	US-09-849-031A-2	Sequence 2, Appli
124	41.5	25.0	730	9	US-09-978-192A-102	Sequence 102, App	197	39	23.5	370	10	US-09-849-562A-2	Sequence 2, Appli
125	41.5	25.0	730	9	US-09-999-832A-102	Sequence 102, App	198	39	23.5	376	10	US-09-818-512-2	Sequence 2, Appli
126	41.5	25.0	730	9	US-09-978-188-102	Sequence 102, App	199	39	23.5	444	10	US-09-881-752A-212	Sequence 212, App
127	41.5	25.0	349	10	US-09-916-031-1	Sequence 1, Appli	200	39	23.5	477	9	US-09-738-626-7048	Sequence 5048, Ap
128	41.5	25.0	378	9	US-09-971-536-46	Sequence 46, Appli	201	39	23.5	503	9	US-10-003-132-4	Sequence 4, Appli
129	41.5	25.0	421	9	US-09-765-272-58	Sequence 58, Appli	202	39	23.5	503	10	US-09-823-038A-51	Sequence 51, Appli
130	41.5	25.0	429	9	US-09-759-130B-76	Sequence 76, Appli	203	39	23.5	547	10	US-09-895-860-2	Sequence 2, Appli
131	41.5	25.0	429	9	US-09-922-683-9	Sequence 9, Appli	204	39	23.5	551	10	US-09-802-472B-21	Sequence 21, Appli
132	41.5	25.0	429	9	US-09-738-626-5542	Sequence 5542, Ap	205	39	23.5	571	9	US-10-036-041-23	Sequence 23, Appli
133	41.5	25.0	511	10	US-09-925-301-1060	Sequence 1060, Ap	206	39	23.5	571	9	US-10-028-072-542	Sequence 542, App
134	41.5	25.0	530	10	US-09-815-242-10105	Sequence 10105, A	207	39	23.5	571	9	US-10-035-855-23	Sequence 23, Appli
135	41.5	25.0	531	10	US-09-815-242-13999	Sequence 13999, A	208	39	23.5	571	9	US-10-121-049-542	Sequence 542, App
136	41.5	25.0	540	10	US-09-729-454-32	Sequence 32, Appli	209	39	23.5	571	9	US-10-123-304-542	Sequence 542, App
137	41.5	25.0	681	9	US-09-759-130B-75	Sequence 75, Appli	210	39	23.5	571	9	US-10-140-470-542	Sequence 542, App
138	41.5	25.0	715	9	US-10-003-132-2	Sequence 2, Appli	211	39	23.5	571	9	US-09-931-836-23	Sequence 23, Appli
139	41.5	25.0	715	9	US-09-759-130B-73	Sequence 73, Appli	212	39	23.5	571	9	US-10-175-346-542	Sequence 542, App
140	41.5	25.0	906	10	US-09-815-242-10567	Sequence 10567, A	213	39	23.5	571	9	US-10-176-518-542	Sequence 542, App
141	40.5	24.4	624	9	US-09-738-626-4289	Sequence 4289, A	214	39	23.5	571	9	US-10-176-521-542	Sequence 542, App
142	40.5	24.4	5215	9	US-09-860-846-2	Sequence 2, Appli	215	39	23.5	571	9	US-10-227-884-210	Sequence 210, App
143	40.5	24.4	5215	10	US-09-861-289-2	Sequence 2, Appli	216	39	23.5	571	12	US-10-036-342-23	Sequence 23, Appli
144	40.5	24.4	17	9	US-09-996-634-90	Sequence 90, Appli	217	39	23.5	572	10	US-09-828-302-11	Sequence 11, Appli
145	40.5	24.4	246	9	US-09-909-567B-49	Sequence 49, Appli	218	39	23.5	711	10	US-09-765-272-82	Sequence 82, Appli
146	40.5	24.4	249	9	US-10-136-517-17	Sequence 17, Appli	219	39	23.5	777	9	US-10-174-590-348	Sequence 348, App
147	40.5	24.4	263	10	US-09-888-623-3	Sequence 3, Appli	220	39	23.5	777	9	US-10-174-590-348	Sequence 348, App
148	40.5	24.4	377	9	US-09-738-626-3873	Sequence 3873, Ap	221	39	23.5	777	9	US-10-176-758-348	Sequence 348, App
149	40.5	24.4	403	9	US-10-034-158-3	Sequence 3, Appli	222	39	23.5	777	9	US-10-175-737-348	Sequence 348, App
150	40.5	24.4	403	10	US-09-086-118-23	Sequence 23, Appli	223	39	23.5	777	9	US-10-173-706-348	Sequence 348, App
151	40.5	24.4	403	10	US-09-835-684-3	Sequence 3, Appli	224	39	23.5	777	9	US-10-175-738-348	Sequence 348, App
152	40.5	24.4	403	10	US-09-880-371-3	Sequence 3, Appli	225	39	23.5	777	9	US-10-175-738-348	Sequence 348, App
153	40.5	24.4	403	10	US-09-879-248-3	Sequence 3, Appli	226	39	23.5	777	9	US-10-176-482-348	Sequence 348, App
154	40.5	24.4	403	10	US-09-770-693-3	Sequence 3, Appli	227	39	23.5	777	9	US-10-176-482-348	Sequence 348, App
155	40.5	24.4	403	10	US-09-766-348-3	Sequence 3, Appli	228	39	23.5	777	9	US-10-176-482-348	Sequence 348, App
156	40.5	24.4	423	10	US-09-073-009-142	Sequence 142, App	229	39	23.5	777	9	US-10-180-552-348	Sequence 348, App
157	40.5	24.4	423	10	US-09-793-306-142	Sequence 142, App	230	39	23.5	777	9	US-10-173-700-348	Sequence 348, App
158	40.5	24.4	488	9	US-09-996-634-140	Sequence 140, App	231	39	23.5	777	9	US-10-174-720-348	Sequence 348, App
159	40.5	24.4	513	9	US-09-738-626-4772	Sequence 4772, Ap	232	39	23.5	777	9	US-10-174-720-348	Sequence 348, App
160	40.5	24.4	548	9	US-09-738-626-3855	Sequence 3855, Ap	233	39	23.5	777	9	US-10-174-582-348	Sequence 348, App
161	40.5	24.4	710	10	US-09-287-849-16	Sequence 16, Appli	234	39	23.5	777	9	US-10-174-588-348	Sequence 348, App
162	40.5	24.4	788	10	US-09-391-340-6	Sequence 6, Appli	235	39	23.5	777	9	US-10-175-739-348	Sequence 348, App
163	40.5	24.4	788	10	US-09-948-369-6	Sequence 6, Appli	236	39	23.5	777	9	US-10-175-740-348	Sequence 348, App
164	40.5	24.4	856	10	US-09-287-849-12	Sequence 12, Appli	237	39	23.5	777	9	US-10-175-743-348	Sequence 348, App
165	40.5	24.4	940	9	US-09-811-088-11	Sequence 11, Appli	238	39	23.5	777	9	US-10-176-488-348	Sequence 348, App

229	39	23.5	777	9	US-10-176-492-348	Sequence 348, App	312	38	22.9	467	9	US-10-176-513-16	Sequence 16, Appl
240	39	23.5	777	9	US-10-176-747-348	Sequence 348, App	313	38	22.9	467	9	US-10-180-552-16	Sequence 16, Appl
241	39	23.5	777	9	US-10-176-750-348	Sequence 348, App	314	38	22.9	467	9	US-10-180-557-16	Sequence 16, Appl
242	39	23.5	777	9	US-10-176-985-348	Sequence 348, App	315	38	22.9	467	9	US-09-148-545-134	Sequence 134, App
243	39	23.5	777	9	US-10-176-987-348	Sequence 348, App	316	38	22.9	467	9	US-09-148-548-195	Sequence 195, App
244	39	23.5	777	9	US-10-176-991-348	Sequence 348, App	317	38	22.9	467	9	US-09-907-613-195	Sequence 195, App
245	39	23.5	777	9	US-10-176-992-348	Sequence 348, App	318	38	22.9	467	9	US-09-907-642-195	Sequence 195, App
246	39	23.5	777	9	US-10-176-993-348	Sequence 348, App	319	38	22.9	467	9	US-10-173-700-16	Sequence 16, Appl
247	39	23.5	777	9	US-10-184-658-348	Sequence 348, App	320	38	22.9	467	9	US-10-174-572-16	Sequence 16, Appl
248	39	23.5	777	12	US-10-052-586-348	Sequence 348, App	321	38	22.9	467	9	US-10-174-579-16	Sequence 16, Appl
249	39	23.5	832	10	US-09-729-674-176	Sequence 176, App	322	38	22.9	467	9	US-10-174-582-16	Sequence 16, Appl
250	39	23.5	966	10	US-09-771-1616-157	Sequence 197, App	323	38	22.9	467	9	US-10-174-588-16	Sequence 16, Appl
251	38.5	23.2	102	10	US-09-864-761-37086	Sequence 37086, A	324	38	22.9	467	9	US-10-175-730-16	Sequence 16, Appl
252	38.5	23.2	233	9	US-09-975-719-81	Sequence 81, Appl	325	38	22.9	467	9	US-10-175-740-16	Sequence 16, Appl
253	38.5	23.2	273	9	US-10-063-547-168	Sequence 168, Appl	326	38	22.9	467	9	US-10-175-743-16	Sequence 16, Appl
254	38.5	23.2	273	9	US-10-036-041-18	Sequence 50, App	327	38	22.9	467	9	US-10-176-488-16	Sequence 16, Appl
255	38.5	23.2	273	9	US-10-028-072-540	Sequence 50, App	328	38	22.9	467	9	US-10-176-492-16	Sequence 16, Appl
256	38.5	23.2	273	9	US-10-035-855-18	Sequence 18, Appl	329	38	22.9	467	9	US-10-176-492-16	Sequence 16, Appl
257	38.5	23.2	273	9	US-10-063-616-168	Sequence 168, App	330	38	22.9	467	9	US-10-176-747-16	Sequence 16, Appl
258	38.5	23.2	273	9	US-10-063-502-168	Sequence 168, App	331	38	22.9	467	9	US-10-176-750-16	Sequence 16, Appl
259	38.5	23.2	273	9	US-10-121-049-540	Sequence 540, App	332	38	22.9	467	9	US-10-176-985-16	Sequence 16, Appl
260	38.5	23.2	273	9	US-10-123-904-540	Sequence 540, App	333	38	22.9	467	9	US-10-176-987-16	Sequence 16, Appl
261	38.5	23.2	273	9	US-10-133-904-540	Sequence 540, App	334	38	22.9	467	9	US-10-176-991-16	Sequence 16, Appl
262	38.5	23.2	273	9	US-10-140-470-540	Sequence 540, App	335	38	22.9	467	9	US-10-176-992-16	Sequence 16, Appl
263	38.5	23.2	273	9	US-09-931-836-18	Sequence 18, Appl	336	38	22.9	467	9	US-10-176-993-16	Sequence 16, Appl
264	38.5	23.2	273	9	US-10-175-746-440	Sequence 540, App	337	38	22.9	467	9	US-10-184-658-16	Sequence 16, Appl
265	38.5	23.2	273	9	US-10-176-918-540	Sequence 540, App	338	38	22.9	467	9	US-09-909-320-195	Sequence 195, App
266	38.5	23.2	273	9	US-10-176-921-540	Sequence 540, App	339	38	22.9	467	12	US-10-052-586-16	Sequence 16, Appl
267	38.5	23.2	273	12	US-10-036-342-18	Sequence 18, Appl	340	38	22.9	467	9	US-09-909-320-195	Sequence 195, App
268	38.5	23.2	273	12	US-10-006-867-168	Sequence 168, Appl	341	38	22.9	467	9	US-10-052-586-16	Sequence 16, Appl
269	38.5	23.2	1150	10	US-10-029-180-44	Sequence 44, Appl	342	38	22.9	467	9	US-10-016-283-36	Sequence 36, Appl
270	38	22.9	26	10	US-09-946-239-9	Sequence 9, Appl	343	38	22.9	467	10	US-09-843-845-16	Sequence 16, Appl
271	38	22.9	66	10	US-09-864-761-41030	Sequence 41030, A	344	38	22.9	520	10	US-09-765-272-122	Sequence 12, Appl
272	38	22.9	66	10	US-09-872-523-66	Sequence 66, Appl	345	38	22.9	521	10	US-09-815-242-11855	Sequence 11855, A
273	38	22.9	66	10	US-09-925-300-1832	Sequence 1832, Ap	346	38	22.9	534	9	US-09-738-626-5899	Sequence 5899, Ap
274	38	22.9	81	9	US-09-894-882-74	Sequence 74, Appl	347	38	22.9	539	10	US-09-945-826-10	Sequence 10, Appl
275	38	22.9	81	9	US-09-861-097-45	Sequence 45, Appl	348	38	22.9	542	10	US-09-805-456A-4	Sequence 4, Appl
276	38	22.9	162	9	US-09-864-761-43917	Sequence 43917, A	349	38	22.9	549	10	US-09-805-456A-2	Sequence 2, Appl
277	38	22.9	177	9	US-09-796-682-2342	Sequence 2342, Ap	350	38	22.9	658	10	US-09-815-242-11824	Sequence 11824, A
278	38	22.9	215	10	US-10-105-427-2	Sequence 2, Appl	351	38	22.9	788	9	US-10-029-180-92	Sequence 92, Appl
279	38	22.9	216	9	US-09-912-020-362	Sequence 362, App	352	38	22.9	803	10	US-09-968-436B-2	Sequence 2, Appl
280	38	22.9	256	9	US-10-016-283-31	Sequence 31, Appl	353	38	22.9	803	10	US-09-759-010-7	Sequence 7, Appl
281	38	22.9	294	9	US-10-016-283-30	Sequence 30, Appl	354	38	22.9	848	9	US-09-875-353-5	Sequence 5, Appl
282	38	22.9	316	10	US-09-815-242-13929	Sequence 13929, A	355	38	22.9	963	10	US-09-886-466-22	Sequence 22, Appl
283	38	22.9	338	9	US-10-016-283-28	Sequence 28, Appl	356	38	22.9	1037	10	US-09-728-772-55	Sequence 55, Appl
284	38	22.9	338	9	US-10-034-158-1	Sequence 1, Appl	357	38	22.9	1040	10	US-09-765-272-118	Sequence 118, App
285	38	22.9	338	10	US-09-866-118-21	Sequence 21, Appl	358	38	22.9	1167	10	US-09-815-242-11522	Sequence 11522, A
286	38	22.9	338	10	US-09-835-684-1	Sequence 1, Appl	359	38	22.9	1169	9	US-09-738-626-4140	Sequence 4140, Ap
287	38	22.9	338	10	US-09-880-371-1	Sequence 1, Appl	360	38	22.6	2905	8	US-08-424-950B-401	Sequence 401, App
288	38	22.9	338	10	US-09-879-248-1	Sequence 1, Appl	361	37.5	22.6	50	10	US-09-764-868-899	Sequence 899, App
289	38	22.9	338	10	US-09-770-693-1	Sequence 1, Appl	362	37.5	22.6	76	10	US-09-864-761-36015	Sequence 36015, A
290	38	22.9	338	10	US-09-766-348-1	Sequence 1, Appl	363	37.5	22.6	105	10	US-09-925-299-997	Sequence 927, App
291	38	22.9	362	9	US-09-766-348-12	Sequence 12, App	364	37.5	22.6	105	10	US-09-815-242-11799	Sequence 11799, A
292	38	22.9	390	9	US-10-016-283-47	Sequence 47, Appl	365	37.5	22.6	105	10	US-09-741-668-346	Sequence 346, App
293	38	22.9	419	10	US-09-740-033-4	Sequence 4, Appl	366	37.5	22.6	475	9	US-09-764-868-899	Sequence 899, App
294	38	22.9	440	9	US-10-016-283-46	Sequence 26, Appl	367	37.5	22.6	640	10	US-09-815-242-11884	Sequence 11884, A
295	38	22.9	456	9	US-10-016-283-25	Sequence 26, Appl	368	37.5	22.6	848	10	US-09-839-185-8	Sequence 8, Appl
296	38	22.9	463	9	US-09-738-626-6946	Sequence 6946, Ap	369	37.5	22.6	1337	10	US-09-803-126-1	Sequence 1, Appl
297	38	22.9	467	9	US-09-905-291A-195	Sequence 195, App	370	37	22.3	19	10	US-09-864-761-46551	Sequence 46551, A
298	38	22.9	467	9	US-09-981-876-134	Sequence 134, App	371	37	22.3	33	10	US-09-864-761-46594	Sequence 46594, A
299	38	22.9	467	9	US-09-902-853-195	Sequence 195, App	372	37	22.3	34	10	US-09-917-265-21	Sequence 21, Appl
300	38	22.9	467	9	US-09-907-824-195	Sequence 195, App	373	37	22.3	129	9	US-09-726-489-10	Sequence 10, Appl
301	38	22.9	467	9	US-09-907-841-195	Sequence 195, App	374	37	22.3	152	9	US-10-092-243A-17	Sequence 17, Appl
302	38	22.9	467	9	US-09-904-011-195	Sequence 195, App	375	37	22.3	155	9	US-10-108-605-361	Sequence 361, App
303	38	22.9	467	9	US-10-174-590-16	Sequence 16, Appl	376	37	22.3	112	10	US-09-815-242-5069	Sequence 5069, App
304	38	22.9	467	9	US-10-176-758-16	Sequence 16, Appl	377	37	22.3	247	9	US-09-738-626-6345	Sequence 6345, Ap
305	38	22.9	467	9	US-10-175-737-16	Sequence 16, Appl	378	37	22.3	302	10	US-09-864-761-35790	Sequence 35790, A
306	38	22.9	467	9	US-09-906-742-195	Sequence 195, App	379	37	22.3	315	10	US-09-886-056-307	Sequence 307, App
307	38	22.9	467	9	US-10-173-706-16	Sequence 16, Appl	380	37	22.3	340	9	US-10-155-947-2	Sequence 2, Appl
308	38	22.9	467	9	US-10-173-706-16	Sequence 16, Appl	381	37	22.3	340	9	US-10-155-947-2	Sequence 2, Appl
309	38	22.9	467	9	US-10-175-752-16	Sequence 16, Appl	382	37	22.3	340	9	US-10-155-947-6	Sequence 6, Appl
310	38	22.9	467	9	US-10-176-482-16	Sequence 16, Appl	383	37	22.3	340	10	US-09-775-195-2	Sequence 2, Appl
311	38	22.9	467	9	US-10-176-757-16	Sequence 16, Appl	384	37	22.3	340	10	US-09-775-195-5	Sequence 5, Appl

385	37	22.3	340	10	US-09-775-195-6	Sequence 6, Appl1	458	22.0	832	9	US-10-033-297-4	Sequence 4, Appl1
386	37	22.3	342	10	US-09-815-242-11043	Sequence 11043, A	459	36.5	832	9	US-09-972-834-2	Sequence 2, Appl1
387	37	22.3	345	9	US-10-155-947-4	Sequence 4, Appl1	460	22.0	832	9	US-10-071-505-1	Sequence 1, Appl1
388	37	22.3	346	10	US-09-775-195-4	Sequence 2, Appl1	461	36.5	832	9	US-10-081-806-4	Sequence 4, Appl1
389	37	22.3	345	10	US-09-829-482-2	Sequence 17, Appl1	462	22.0	833	9	US-10-074-328-4	Sequence 8, Appl1
390	37	22.3	366	10	US-09-901-252-17	Sequence 1020, Ap	463	36.5	833	9	US-10-033-297-6	Sequence 6, Appl1
391	37	22.3	374	9	US-09-764-868-1020	Sequence 50, Appl	464	22.0	833	9	US-10-033-297-9	Sequence 6, Appl1
392	37	22.3	379	10	US-09-934-868-50	Sequence 4635, Ap	465	36.5	833	9	US-10-033-297-11	Sequence 7, Appl1
393	37	22.3	392	9	US-09-738-626-4635	Sequence 15, Appl	466	22.0	833	9	US-10-033-297-7	Sequence 8, Appl1
394	37	22.3	410	10	US-09-917-265-15	Sequence 24, Appl	467	36.5	833	9	US-10-081-806-8	Sequence 8, Appl1
395	37	22.3	416	10	US-09-917-265-24	Sequence 6944, Ap	468	22.0	833	9	US-10-074-328-8	Sequence 8, Appl1
396	37	22.3	426	9	US-09-738-626-6944	Sequence 251, App	469	36.5	1016	12	US-10-007-693-95	Sequence 95, Appl1
397	37	22.3	438	9	US-09-712-363-251	Sequence 3703, Ap	470	21.7	24	10	US-09-864-761-47163	Sequence 47163, A
398	37	22.3	467	10	US-09-863-547B-1	Sequence 313, App	471	36	35	10	US-09-864-882-45369	Sequence 45369, A
399	37	22.3	491	10	US-09-738-626-3703	Sequence 10504, A	472	36	68	10	US-09-925-442-18	Sequence 18, Appl
400	37	22.3	532	10	US-09-820-658-313	Sequence 13812, A	473	36	63	10	US-09-864-761-48204	Sequence 48204, A
401	37	22.3	648	10	US-09-815-242-10504	Sequence 5453, Ap	474	36	82	10	US-09-764-870-11067	Sequence 11067, Ap
402	37	22.3	654	10	US-09-815-242-13812	Sequence 18, Appl	475	36	96	10	US-09-833-067-11	Sequence 11, Appl
403	37	22.3	722	9	US-09-738-626-5453	Sequence 5, Appl	476	36	102	10	US-09-925-299-783	Sequence 783, App
404	37	22.3	732	9	US-09-921-496-18	Sequence 18, Appl	477	36	125	12	US-10-078-929-158	Sequence 158, App
405	37	22.3	732	10	US-09-739-010-5	Sequence 5, Appl	478	36	149	10	US-09-767-041-31	Sequence 31, Appl
406	37	22.3	732	10	US-09-874-923-18	Sequence 4131, Ap	479	36	152	10	US-09-939-980-484	Sequence 484, App
407	37	22.3	799	9	US-09-738-626-4131	Sequence 3, Appl1	480	36	158	10	US-09-764-870-377	Sequence 377, App
408	37	22.3	802	10	US-09-758-386-3	Sequence 113, App	481	36	158	10	US-09-764-860-342	Sequence 342, App
409	37	22.3	805	9	US-10-108-605-113	Sequence 5418, Ap	482	36	178	9	US-10-076-785-30	Sequence 30, Appl
410	37	22.3	1100	10	US-09-815-242-11861	Sequence 5116, Ap	483	36	195	10	US-09-833-067-2	Sequence 2, Appl1
411	37	22.3	1283	9	US-09-738-626-5418	Sequence 6, Appl1	484	36	195	10	US-09-833-067-6	Sequence 6, Appl1
412	37	22.3	1317	10	US-09-815-242-5118	Sequence 3, Appl1	485	36	195	10	US-09-833-067-8	Sequence 8, Appl1
413	37	22.3	1430	10	US-09-740-274-6	Sequence 3, Appl1	486	36	195	10	US-09-833-067-32	Sequence 32, Appl1
414	37	22.3	2037	9	US-09-951-402-3	Sequence 14, Appl	487	36	195	10	US-09-833-079-36	Sequence 36, Appl
415	37	22.3	2037	10	US-09-923-101-3	Sequence 33, Appl	488	36	195	10	US-09-833-079-39	Sequence 39, Appl
416	37	22.3	2224	9	US-10-115-563-14	Sequence 6, Appl1	489	36	195	10	US-09-978-2854-477	Sequence 477, App
417	37	22.3	3729	9	US-09-861-289-33	Sequence 37259, A	490	36	201	9	US-09-978-1924-477	Sequence 477, App
418	37	22.3	3739	10	US-09-861-289-33	Sequence 35614, A	491	36	201	9	US-09-999-8324-477	Sequence 477, App
419	37	22.3	3759	10	US-09-861-289-33	Sequence 1608, Ap	492	36	201	9	US-09-978-959-477	Sequence 477, App
420	37	22.3	11877	9	US-09-861-289-6	Sequence 1417, Ap	493	36	201	9	US-10-174-580-108	Sequence 108, App
421	37	22.3	11877	10	US-09-861-289-6	Sequence 1885, Ap	494	36	201	9	US-10-176-757-108	Sequence 108, App
422	36.5	22.0	55	10	US-09-864-761-37259	Sequence 669, App	495	36	201	9	US-10-175-727-108	Sequence 108, App
423	36.5	22.0	62	10	US-09-864-761-35614	Sequence 4931, Ap	496	36	201	9	US-10-173-706-108	Sequence 108, App
424	36.5	22.0	65	10	US-09-923-300-1608	Sequence 1175, Ap	497	36	201	9	US-10-175-728-108	Sequence 108, App
425	36.5	22.0	77	9	US-09-796-692-1417	Sequence 247, App	498	36	201	9	US-10-176-482-108	Sequence 108, App
426	36.5	22.0	77	9	US-09-796-692-1885	Sequence 620, App	499	36	201	9	US-10-176-757-108	Sequence 108, App
427	36.5	22.0	79	9	US-09-796-692-669	Sequence 61, Appl1	500	36	201	9	US-10-176-757-108	Sequence 108, App
428	36.5	22.0	101	9	US-09-738-626-4931	Sequence 2, Appl1	501	36	201	9	US-10-176-757-108	Sequence 108, App
429	36.5	22.0	105	10	US-09-764-869-1175	Sequence 1, Appl1	502	36	201	9	US-10-176-757-108	Sequence 108, App
430	36.5	22.0	240	9	US-09-989-920-247	Sequence 3, Appl1	503	36	201	9	US-10-176-757-108	Sequence 108, App
431	36.5	22.0	305	10	US-09-925-302-620	Sequence 268, App	504	36	201	9	US-10-176-757-108	Sequence 108, App
432	36.5	22.0	320	9	US-10-033-297-61	Sequence 13180, A	505	36	201	9	US-10-176-757-108	Sequence 108, App
433	36.5	22.0	339	10	US-09-943-671-2	Sequence 10200, A	506	36	201	9	US-10-176-757-108	Sequence 108, App
434	36.5	22.0	339	12	US-10-067-291-1	Sequence 13621, A	507	36	201	9	US-10-176-757-108	Sequence 108, App
435	36.5	22.0	377	10	US-09-849-562A-3	Sequence 6778, Ap	508	36	201	9	US-10-176-757-108	Sequence 108, App
436	36.5	22.0	377	10	US-09-849-562A-3	Sequence 6732, Ap	509	36	201	9	US-10-176-757-108	Sequence 108, App
437	36.5	22.0	408	9	US-09-712-363-268	Sequence 16, Appl	510	36	201	9	US-10-176-757-108	Sequence 108, App
438	36.5	22.0	431	10	US-09-815-242-13180	Sequence 813, App	511	36	201	9	US-10-176-757-108	Sequence 108, App
439	36.5	22.0	431	10	US-09-815-242-13639	Sequence 18, Appl	512	36	201	9	US-10-176-757-108	Sequence 108, App
440	36.5	22.0	475	10	US-09-815-242-10200	Sequence 31, Appl	513	36	201	9	US-10-176-757-108	Sequence 108, App
441	36.5	22.0	476	10	US-09-815-242-13821	Sequence 329, App	514	36	201	9	US-10-176-757-108	Sequence 108, App
442	36.5	22.0	498	9	US-09-738-626-6678	Sequence 3, Appl1	515	36	201	9	US-10-176-757-108	Sequence 108, App
443	36.5	22.0	507	9	US-09-738-626-6732	Sequence 16, Appl	516	36	201	9	US-10-176-757-108	Sequence 108, App
444	36.5	22.0	525	10	US-09-731-393-16	Sequence 8, Appl1	517	36	201	9	US-10-176-757-108	Sequence 108, App
445	36.5	22.0	527	10	US-09-731-393-10	Sequence 7, Appl1	518	36	201	9	US-10-176-757-108	Sequence 108, App
446	36.5	22.0	535	10	US-09-815-242-5557	Sequence 7, Appl1	519	36	201	9	US-10-176-757-108	Sequence 108, App
447	36.5	22.0	540	10	US-09-815-242-12510	Sequence 18, Appl	520	36	201	9	US-10-176-757-108	Sequence 108, App
448	36.5	22.0	545	10	US-09-988-200-8	Sequence 31, Appl	521	36	201	9	US-10-176-757-108	Sequence 108, App
449	36.5	22.0	566	10	US-09-943-671-7	Sequence 329, App	522	36	201	9	US-10-176-757-108	Sequence 108, App
450	36.5	22.0	613	10	US-09-943-671-7	Sequence 3, Appl1	523	36	201	9	US-10-176-757-108	Sequence 108, App
451	36.5	22.0	655	9	US-09-949-842-18	Sequence 18, Appl	524	36	201	9	US-10-176-757-108	Sequence 108, App
452	36.5	22.0	655	10	US-09-943-671-31	Sequence 31, Appl	525	36	201	9	US-10-176-757-108	Sequence 108, App
453	36.5	22.0	715	10	US-09-841-132-329	Sequence 329, App	526	36	201	9	US-10-176-757-108	Sequence 108, App
454	36.5	22.0	830	9	US-10-071-503-3	Sequence 3, Appl1	527	36	201	9	US-10-176-757-108	Sequence 108, App
455	36.5	22.0	831	9	US-10-033-297-5	Sequence 5, Appl1	528	36	201	12	US-10-052-586-1034	Sequence 1034, App
456	36.5	22.0	831	9	US-10-081-806-5	Sequence 5, Appl1	529	36	239	9	US-10-076-785-36	Sequence 36, Appl
457	36.5	22.0	831	9	US-10-074-328-5	Sequence 5, Appl1	530	36	244	9	US-10-028-072-156	Sequence 156, App

531	36	21.7	244	9	US-10-121-049-156	Sequence 156, App	604	36	21.7	865	9	US-10-055-364-24	Sequence 24, Appl
532	36	21.7	244	9	US-10-123-904-156	Sequence 156, App	605	36	21.7	966	10	US-09-815-242-5251	Sequence 5251, Ap
533	36	21.7	244	9	US-10-140-470-156	Sequence 156, App	606	36	21.7	1008	12	US-10-017-828-7	Sequence 7, Appl1
534	36	21.7	244	9	US-10-175-746-156	Sequence 156, App	607	36	21.7	1008	12	US-10-017-828-8	Sequence 8, Appl1
535	36	21.7	244	9	US-10-16-918-156	Sequence 156, App	608	36	21.7	1009	10	US-09-815-242-12141	Sequence 12141, A
536	36	21.7	244	9	US-10-176-921-156	Sequence 156, App	609	36	21.7	1193	12	US-10-040-919-2	Sequence 2, Appl1
537	36	21.7	244	10	US-09-731-872-359	Sequence 359, App	610	36	21.7	1239	9	US-10-076-622-577	Sequence 577, App
538	36	21.7	247	9	US-10-076-622-571	Sequence 571, App	611	36	21.7	1239	12	US-10-007-805-577	Sequence 577, App
539	36	21.7	247	12	US-10-007-805-571	Sequence 571, App	612	36	21.7	1341	9	US-10-076-622-565	Sequence 565, App
540	36	21.7	258	10	US-09-815-242-4947	Sequence 4947, Ap	613	36	21.7	1341	12	US-10-007-805-565	Sequence 565, App
541	36	21.7	258	10	US-09-840-704-3	Sequence 3, Appl1	614	36	21.7	1341	12	US-10-076-622-573	Sequence 573, App
542	36	21.7	261	10	US-09-815-242-10487	Sequence 10487, A	615	36	21.7	1349	12	US-10-007-805-573	Sequence 573, App
543	36	21.7	283	9	US-09-738-626-3603	Sequence 3603, Ap	616	36	21.7	1642	10	US-09-925-442-2	Sequence 2, Appl1
544	36	21.7	288	12	US-10-062-254-286	Sequence 286, App	617	36	21.7	1648	10	US-09-925-442-35	Sequence 35, Appl1
545	36	21.7	290	10	US-09-881-752A-270	Sequence 270, App	618	36	21.7	2697	10	US-09-961-557A-5	Sequence 5, Appl1
546	36	21.7	292	10	US-09-886-055-315	Sequence 315, App	619	36	21.7	2972	8	US-08-424-550B-187	Sequence 387, App
547	36	21.7	293	9	US-09-971-536-71	Sequence 71, Appl	620	35.5	21.4	38	10	US-09-925-301-1495	Sequence 1495, Ap
548	36	21.7	298	10	US-09-801-368-66	Sequence 66, Appl	621	35.5	21.4	8	10	US-09-925-300-1746	Sequence 1746, Ap
549	36	21.7	303	9	US-09-854-133-438	Sequence 438, App	622	35.5	21.4	8	10	US-09-764-847-934	Sequence 934, App
550	36	21.7	303	10	US-09-738-973-438	Sequence 5340, Ap	623	35.5	21.4	86	10	US-09-764-860-544	Sequence 544, App
551	36	21.7	321	9	US-09-738-626-5340	Sequence 1, Appl1	624	35.5	21.4	126	10	US-09-815-242-5124	Sequence 5124, Ap
552	36	21.7	336	10	US-09-747-804-1	Sequence 36, Appl	625	35.5	21.4	187	9	US-09-992-598-189	Sequence 189, App
553	36	21.7	339	10	US-09-970-711-36	Sequence 9, Appl1	626	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
554	36	21.7	367	9	US-09-939-633-9	Sequence 9, Appl1	627	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
555	36	21.7	367	10	US-09-939-754-9	Sequence 9, Appl1	628	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
556	36	21.7	367	10	US-09-839-832-9	Sequence 9, Appl1	629	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
557	36	21.7	378	10	US-09-815-242-5137	Sequence 5137, Ap	630	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
558	36	21.7	384	9	US-09-924-400-326	Sequence 326, App	631	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
559	36	21.7	384	9	US-09-924-400-332	Sequence 332, App	632	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
560	36	21.7	384	9	US-09-924-400-334	Sequence 334, App	633	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
561	36	21.7	384	10	US-09-810-936-326	Sequence 326, App	634	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
562	36	21.7	384	10	US-09-810-936-332	Sequence 332, App	635	35.5	21.4	187	9	US-09-989-733-189	Sequence 189, App
563	36	21.7	384	9	US-09-810-936-334	Sequence 334, App	636	35.5	21.4	187	9	US-10-176-590-426	Sequence 426, App
564	36	21.7	382	9	US-10-094-080-1	Sequence 1, Appl1	637	35.5	21.4	187	9	US-10-176-590-426	Sequence 426, App
565	36	21.7	395	9	US-09-712-563-164	Sequence 164, App	638	35.5	21.4	187	9	US-10-176-590-426	Sequence 426, App
566	36	21.7	397	10	US-09-842-987-2	Sequence 2, Appl1	639	35.5	21.4	187	9	US-09-993-667-189	Sequence 189, App
567	36	21.7	399	9	US-10-076-622-570	Sequence 570, App	640	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
568	36	21.7	399	12	US-10-076-622-572	Sequence 572, App	641	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
569	36	21.7	399	12	US-10-007-805-570	Sequence 570, App	642	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
570	36	21.7	399	12	US-10-007-805-572	Sequence 572, App	643	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
571	36	21.7	426	12	US-10-078-929-160	Sequence 160, App	644	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
572	36	21.7	450	9	US-09-977-660-7	Sequence 7, Appl1	645	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
573	36	21.7	450	10	US-09-815-242-5214	Sequence 5214, Ap	646	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
574	36	21.7	450	10	US-09-977-269-7	Sequence 7, Appl1	647	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
575	36	21.7	523	10	US-09-912-020-354	Sequence 354, App	648	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
576	36	21.7	536	10	US-09-741-660-465	Sequence 465, App	649	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
577	36	21.7	550	9	US-10-076-185-32	Sequence 32, Appl	650	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
578	36	21.7	550	9	US-09-738-626-5843	Sequence 5843, Ap	651	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
579	36	21.7	550	10	US-09-943-671-25	Sequence 25, Appl	652	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
580	36	21.7	557	9	US-09-778-319-2	Sequence 2, Appl1	653	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
581	36	21.7	605	9	US-10-028-072-160	Sequence 160, App	654	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
582	36	21.7	605	9	US-10-121-049-160	Sequence 160, App	655	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
583	36	21.7	605	9	US-10-123-904-160	Sequence 160, App	656	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
584	36	21.7	605	9	US-10-140-470-160	Sequence 160, App	657	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
585	36	21.7	605	9	US-09-832-129-37	Sequence 37, Appl	658	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
586	36	21.7	605	9	US-10-173-746-160	Sequence 160, App	659	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
587	36	21.7	605	9	US-10-176-918-160	Sequence 160, App	660	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
588	36	21.7	605	9	US-10-176-921-160	Sequence 160, App	661	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
589	36	21.7	605	10	US-09-860-298-2	Sequence 2, Appl1	662	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
590	36	21.7	630	9	US-09-738-626-6059	Sequence 6059, Ap	663	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
591	36	21.7	683	10	US-09-770-657-2	Sequence 2, Appl1	664	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
592	36	21.7	683	12	US-10-025-676-1	Sequence 1, Appl1	665	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
593	36	21.7	709	10	US-09-815-242-13349	Sequence 13349, A	666	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
594	36	21.7	740	10	US-09-771-161A-265	Sequence 265, App	667	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
595	36	21.7	740	10	US-09-771-161A-265	Sequence 265, App	668	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
596	36	21.7	745	9	US-09-844-988-10	Sequence 10, Appl	669	35.5	21.4	187	9	US-10-173-706-426	Sequence 426, App
597	36	21.7	745	10	US-09-796-873-2	Sequence 2, Appl1	670	35.5	21.4	187	10	US-09-989-722-189	Sequence 189, App
598	36	21.7	745	10	US-09-844-908-10	Sequence 10, Appl	671	35.5	21.4	187	10	US-09-989-722-189	Sequence 189, App
599	36	21.7	760	9	US-09-712-363-292	Sequence 292, Appl	672	35.5	21.4	187	10	US-09-989-727-189	Sequence 189, App
600	36	21.7	760	9	US-10-076-785-28	Sequence 28, Appl	673	35.5	21.4	187	10	US-09-989-727-189	Sequence 189, App
601	36	21.7	762	10	US-09-804-969-15	Sequence 15, Appl1	674	35.5	21.4	187	10	US-09-989-727-189	Sequence 189, App
602	36	21.7	762	10	US-09-908-664-2	Sequence 2, Appl1	675	35.5	21.4	187	10	US-09-989-727-189	Sequence 189, App
603	36	21.7	792	10	US-09-815-242-11880	Sequence 11880, A	676	35.5	21.4	187	10	US-09-989-727-189	Sequence 189, App

677	35.5	21.4	187	10	US-09-990-442-189	Sequence 189, App	750	35	21.1	166	9	US-10-175-746-8	Sequence 8, Appl1
678	35.5	21.4	187	10	US-09-991-163-189	Sequence 189, App	751	35	21.1	166	9	US-10-176-918-8	Sequence 8, Appl1
679	35.5	21.4	187	10	US-09-993-604-189	Sequence 189, App	752	35	21.1	166	9	US-10-176-921-8	Sequence 8, Appl1
680	35.5	21.4	187	10	US-09-990-456-189	Sequence 189, App	753	35	21.1	171	10	US-09-864-761-43383	Sequence 43383, A
681	35.5	21.4	187	10	US-09-989-721-189	Sequence 189, App	754	35	21.1	176	10	US-09-746-284-12	Sequence 12, Appl1
682	35.5	21.4	187	12	US-10-052-586-426	Sequence 426, App	755	35	21.1	194	10	US-09-772-105-19	Sequence 19, Appl1
683	35.5	21.4	196	10	US-09-925-299-1008	Sequence 1008, App	756	35	21.1	203	9	US-09-738-626-6095	Sequence 6095, App
684	35.5	21.4	252	10	US-09-893-737-52	Sequence 52, Appl1	757	35	21.1	204	9	US-09-949-842-25	Sequence 25, Appl1
685	35.5	21.4	254	10	US-09-815-342-13859	Sequence 13859, A	758	35	21.1	237	10	US-09-925-301-854	Sequence 854, App
686	35.5	21.4	263	9	US-09-924-340-40	Sequence 40, Appl1	759	35	21.1	254	9	US-09-738-626-4681	Sequence 4681, App
687	35.5	21.4	263	9	US-09-992-600A-40	Sequence 40, Appl1	760	35	21.1	264	9	US-09-510-332-48	Sequence 48, Appl1
688	35.5	21.4	266	10	US-09-815-242-10835	Sequence 10835, A	761	35	21.1	268	10	US-09-772-105-12	Sequence 12, Appl1
689	35.5	21.4	290	9	US-09-738-626-4723	Sequence 4723, App	762	35	21.1	285	9	US-09-738-626-5716	Sequence 5716, App
690	35.5	21.4	298	10	US-09-925-297-802	Sequence 802, App	763	35	21.1	286	10	US-09-815-242-13438	Sequence 13438, A
691	35.5	21.4	317	10	US-09-825-561A-84	Sequence 84, Appl1	764	35	21.1	286	10	US-09-815-242-13535	Sequence 13535, A
692	35.5	21.4	321	10	US-09-886-055-259	Sequence 259, App	765	35	21.1	286	10	US-09-943-671-13	Sequence 13, Appl1
693	35.5	21.4	327	10	US-09-870-962-5	Sequence 5, Appl1	766	35	21.1	288	9	US-09-738-626-4503	Sequence 4503, App
694	35.5	21.4	380	9	US-10-104-408-1	Sequence 1, Appl1	767	35	21.1	291	10	US-09-741-650-375	Sequence 375, App
695	35.5	21.4	380	9	US-09-780-926-1	Sequence 1, Appl1	768	35	21.1	297	10	US-09-815-242-11721	Sequence 13721, A
696	35.5	21.4	380	10	US-09-871-617-2	Sequence 2, Appl1	769	35	21.1	303	9	US-09-916-338A-2	Sequence 2, Appl1
697	35.5	21.4	380	10	US-09-764-803A-23	Sequence 23, Appl1	770	35	21.1	303	10	US-09-838-561-13	Sequence 13, Appl1
698	35.5	21.4	404	9	US-09-827-708A-2	Sequence 2, Appl1	771	35	21.1	303	10	US-09-816-750-13	Sequence 13, Appl1
699	35.5	21.4	404	10	US-09-954-697-6	Sequence 6, Appl1	772	35	21.1	328	10	US-09-942-445-4	Sequence 4, Appl1
700	35.5	21.4	404	10	US-09-888-243-30	Sequence 30, Appl1	773	35	21.1	328	10	US-09-795-693-35	Sequence 35, Appl1
701	35.5	21.4	405	9	US-09-738-626-6717	Sequence 6717, App	774	35	21.1	328	10	US-09-938-970-4	Sequence 4, Appl1
702	35.5	21.4	420	10	US-09-946-142-2	Sequence 2, Appl1	775	35	21.1	338	10	US-09-815-242-11750	Sequence 11750, A
703	35.5	21.4	429	9	US-09-854-133-381	Sequence 33, App	776	35	21.1	342	9	US-09-764-854-1146	Sequence 1146, App
704	35.5	21.4	429	10	US-09-835-995-2	Sequence 2, Appl1	777	35	21.1	342	9	US-10-001-857-201	Sequence 201, App
705	35.5	21.4	429	10	US-09-738-626-391	Sequence 391, App	778	35	21.1	354	10	US-09-841-132-576	Sequence 576, App
706	35.5	21.4	438	9	US-09-738-626-6181	Sequence 6181, App	779	35	21.1	369	10	US-09-994-427A-9	Sequence 9, Appl1
707	35.5	21.4	447	9	US-09-843-250-23	Sequence 23, Appl1	780	35	21.1	378	10	US-10-047-260-12	Sequence 12, Appl1
708	35.5	21.4	454	10	US-09-815-242-13445	Sequence 13445, A	781	35	21.1	400	12	US-10-055-430-3	Sequence 3, Appl1
709	35.5	21.4	524	10	US-09-815-242-10942	Sequence 10942, A	782	35	21.1	410	10	US-09-963-896-2	Sequence 2, Appl1
710	35.5	21.4	529	12	US-10-011-436-4	Sequence 4, Appl1	783	35	21.1	436	10	US-09-922-501-8	Sequence 8, Appl1
711	35.5	21.4	578	10	US-09-924-358-39	Sequence 39, Appl1	784	35	21.1	449	9	US-09-738-626-6555	Sequence 6555, App
712	35.5	21.4	621	9	US-10-005-211-2	Sequence 2, Appl1	785	35	21.1	477	10	US-09-815-242-11826	Sequence 11826, A
713	35.5	21.4	709	9	US-09-767-870-8	Sequence 8, Appl1	786	35	21.1	485	10	US-09-815-242-11090	Sequence 11090, A
714	35.5	21.4	729	9	US-10-051-902-29	Sequence 29, Appl1	787	35	21.1	485	10	US-09-748-127-4	Sequence 4, Appl1
715	35.5	21.4	729	10	US-10-051-909-29	Sequence 29, Appl1	788	35	21.1	528	9	US-10-174-590-490	Sequence 490, App
716	35.5	21.4	740	10	US-09-897-214-14	Sequence 14, Appl1	789	35	21.1	528	9	US-10-176-753-490	Sequence 490, App
717	35.5	21.4	816	9	US-09-796-720B-2	Sequence 2, Appl1	790	35	21.1	528	9	US-10-173-706-490	Sequence 490, App
718	35.5	21.4	868	9	US-10-041-007-4	Sequence 4, Appl1	791	35	21.1	528	9	US-10-180-552-490	Sequence 490, App
719	35.5	21.4	868	10	US-09-887-586A-56	Sequence 56, Appl1	792	35	21.1	528	9	US-10-175-738-490	Sequence 490, App
720	35.5	21.4	868	10	US-09-903-012-56	Sequence 17, App	793	35	21.1	528	9	US-10-175-852-490	Sequence 490, App
721	35.5	21.4	1316	9	US-09-712-363-177	Sequence 177, App	794	35	21.1	528	9	US-10-176-757-490	Sequence 490, App
722	35	21.1	26	9	US-09-774-639-302	Sequence 302, App	795	35	21.1	528	9	US-10-176-513-490	Sequence 490, App
723	35	21.1	35	10	US-09-894-882-75	Sequence 75, Appl1	796	35	21.1	528	9	US-10-180-552-490	Sequence 490, App
724	35	21.1	56	10	US-09-864-761-45011	Sequence 45011, A	797	35	21.1	528	9	US-10-180-557-490	Sequence 490, App
725	35	21.1	64	10	US-09-864-761-41283	Sequence 41283, A	798	35	21.1	528	9	US-10-180-557-490	Sequence 490, App
726	35	21.1	70	10	US-09-864-761-46514	Sequence 46514, A	799	35	21.1	528	9	US-09-796-753-122	Sequence 122, App
727	35	21.1	75	10	US-09-864-761-44904	Sequence 44904, A	800	35	21.1	528	9	US-10-173-700-490	Sequence 490, App
728	35	21.1	85	10	US-09-864-761-38557	Sequence 38557, A	801	35	21.1	528	9	US-10-174-572-490	Sequence 490, App
729	35	21.1	85	10	US-09-731-872-374	Sequence 374, App	802	35	21.1	528	9	US-10-174-579-490	Sequence 490, App
730	35	21.1	97	10	US-09-764-869-704	Sequence 704, App	803	35	21.1	528	9	US-10-174-582-490	Sequence 490, App
731	35	21.1	109	10	US-09-910-150-32	Sequence 32, Appl1	804	35	21.1	528	9	US-10-174-588-490	Sequence 490, App
732	35	21.1	121	9	US-09-975-719-421	Sequence 421, App	805	35	21.1	528	9	US-10-175-739-490	Sequence 490, App
733	35	21.1	121	10	US-09-764-898-182	Sequence 182, App	806	35	21.1	528	9	US-10-175-740-490	Sequence 490, App
734	35	21.1	122	9	US-10-001-883-81	Sequence 81, Appl1	807	35	21.1	528	9	US-10-175-743-490	Sequence 490, App
735	35	21.1	139	10	US-09-827-949-31	Sequence 31, Appl1	808	35	21.1	528	9	US-10-176-488-490	Sequence 490, App
736	35	21.1	140	10	US-09-764-877-1654	Sequence 1654, App	809	35	21.1	528	9	US-10-176-492-490	Sequence 490, App
737	35	21.1	142	9	US-09-774-639-298	Sequence 298, App	810	35	21.1	528	9	US-10-176-747-490	Sequence 490, App
738	35	21.1	143	9	US-10-012-896-482	Sequence 482, App	811	35	21.1	528	9	US-10-176-750-490	Sequence 490, App
739	35	21.1	143	9	US-09-895-793-482	Sequence 482, App	812	35	21.1	528	9	US-10-176-885-490	Sequence 490, App
740	35	21.1	143	9	US-09-895-814-482	Sequence 482, App	813	35	21.1	528	9	US-10-176-887-490	Sequence 490, App
741	35	21.1	143	10	US-09-759-143-482	Sequence 482, App	814	35	21.1	528	9	US-10-176-991-490	Sequence 490, App
742	35	21.1	143	10	US-09-780-669-482	Sequence 482, App	815	35	21.1	528	9	US-10-176-992-490	Sequence 490, App
743	35	21.1	143	10	US-09-822-827-482	Sequence 482, App	816	35	21.1	528	9	US-10-176-993-490	Sequence 490, App
744	35	21.1	150	9	US-10-147-874-7	Sequence 7, Appl1	817	35	21.1	528	9	US-10-184-658-490	Sequence 490, App
745	35	21.1	152	9	US-09-738-626-5656	Sequence 5656, App	818	35	21.1	528	10	US-09-732-224-2	Sequence 2, Appl1
746	35	21.1	166	9	US-10-028-072-8	Sequence 8, Appl1	819	35	21.1	528	12	US-10-052-586-490	Sequence 490, App
747	35	21.1	166	9	US-10-121-049-8	Sequence 8, Appl1	820	35	21.1	530	9	US-08-738-626-4780	Sequence 4780, App
748	35	21.1	166	9	US-10-123-904-8	Sequence 8, Appl1	821	35	21.1	532	9	US-09-738-626-6268	Sequence 6268, App
749	35	21.1	166	9	US-10-140-470-8	Sequence 8, Appl1	822	35	21.1	543	10	US-09-732-224-5	Sequence 5, Appl1

823	35	21.1	554	10	US-09-746-491-41	Sequence 41, App1	896	34.5	20.8	150	9	US-09-895-793-707	Sequence 707, App
824	35	21.1	631	9	US-10-106-605-169	Sequence 159, App	897	34.5	20.8	150	9	US-09-895-814-707	Sequence 707, App
825	35	21.1	638	10	US-09-815-242-5222	Sequence 5222, App	898	34.5	20.8	150	10	US-09-759-143-707	Sequence 707, App
826	35	21.1	642	10	US-09-815-242-12143	Sequence 12143, A	899	34.5	20.8	150	10	US-09-780-669-707	Sequence 707, App
827	35	21.1	649	10	US-09-815-242-5333	Sequence 5333, App	900	34.5	20.8	150	10	US-09-822-827-707	Sequence 707, App
828	35	21.1	662	10	US-09-815-242-12343	Sequence 12343, A	901	34.5	20.8	152	10	US-09-994-365-9	Sequence 9, App11
829	35	21.1	663	10	US-09-815-242-10070	Sequence 10070, A	902	34.5	20.8	153	8	US-08-829-402-3	Sequence 3, App11
830	35	21.1	663	10	US-09-815-242-14080	Sequence 14080, A	903	34.5	20.8	158	10	US-09-925-301-1365	Sequence 1365, App
831	35	21.1	692	10	US-09-899-471-8	Sequence 8, App11	904	34.5	20.8	179	9	US-10-021-811-4	Sequence 4, App11
832	35	21.1	705	9	US-09-874-503-14	Sequence 14, App1	905	34.5	20.8	209	10	US-09-815-242-10705	Sequence 12, App
833	35	21.1	705	9	US-10-000-157-14	Sequence 162, App	906	34.5	20.8	249	9	US-09-712-363-172	Sequence 127, App
834	35	21.1	705	9	US-10-063-547-162	Sequence 14, App1	907	34.5	20.8	250	9	US-09-738-626-5662	Sequence 5662, App
835	35	21.1	705	9	US-09-816-744-14	Sequence 14, App1	908	34.5	20.8	280	10	US-09-742-954-8	Sequence 8, App11
836	35	21.1	705	9	US-09-747-359-14	Sequence 58, App	909	34.5	20.8	301	9	US-09-736-626-4278	Sequence 280, App
837	35	21.1	705	9	US-10-174-590-598	Sequence 58, App	910	34.5	20.8	311	10	US-09-731-872-280	Sequence 3, App11
838	35	21.1	705	9	US-10-176-758-598	Sequence 162, App	911	34.5	20.8	376	10	US-09-783-203-3	Sequence 8, App11
839	35	21.1	705	9	US-10-063-516-162	Sequence 58, App	912	34.5	20.8	376	10	US-09-734-357-4	Sequence 10, App11
840	35	21.1	705	9	US-10-175-737-598	Sequence 162, App	913	34.5	20.8	376	10	US-09-734-357-8	Sequence 10, App11
841	35	21.1	705	9	US-10-063-502-162	Sequence 58, App	914	34.5	20.8	376	10	US-09-734-357-10	Sequence 12, App11
842	35	21.1	705	9	US-10-173-706-598	Sequence 58, App	915	34.5	20.8	422	12	US-10-042-417-4	Sequence 12, App11
843	35	21.1	705	9	US-10-175-738-598	Sequence 58, App	916	34.5	20.8	422	12	US-10-042-417-4	Sequence 12, App11
844	35	21.1	705	9	US-10-175-752-598	Sequence 58, App	917	34.5	20.8	423	10	US-09-815-242-14013	Sequence 14, App11
845	35	21.1	705	9	US-10-176-482-598	Sequence 58, App	918	34.5	20.8	451	9	US-09-843-250-24	Sequence 24, App11
846	35	21.1	705	9	US-10-176-757-598	Sequence 58, App	919	34.5	20.8	455	10	US-09-742-954-6	Sequence 6, App11
847	35	21.1	705	9	US-10-176-913-598	Sequence 58, App	920	34.5	20.8	465	9	US-10-108-605-227	Sequence 227, App
848	35	21.1	705	9	US-10-180-552-598	Sequence 58, App	921	34.5	20.8	489	9	US-09-992-598-138	Sequence 138, App
849	35	21.1	705	9	US-10-180-557-598	Sequence 58, App	922	34.5	20.8	489	9	US-09-989-293-138	Sequence 138, App
850	35	21.1	705	9	US-10-173-700-598	Sequence 58, App	923	34.5	20.8	489	9	US-09-989-735-138	Sequence 138, App
851	35	21.1	705	9	US-10-174-572-598	Sequence 58, App	924	34.5	20.8	489	9	US-09-990-444-138	Sequence 138, App
852	35	21.1	705	9	US-10-174-579-598	Sequence 58, App	925	34.5	20.8	489	9	US-09-989-730-138	Sequence 138, App
853	35	21.1	705	9	US-10-174-582-598	Sequence 58, App	926	34.5	20.8	489	9	US-09-990-436-138	Sequence 138, App
854	35	21.1	705	9	US-10-174-588-598	Sequence 58, App	927	34.5	20.8	489	9	US-09-991-181-138	Sequence 138, App
855	35	21.1	705	9	US-10-175-739-598	Sequence 58, App	928	34.5	20.8	489	9	US-09-993-687-138	Sequence 138, App
856	35	21.1	705	9	US-10-175-740-598	Sequence 58, App	929	34.5	20.8	489	9	US-09-989-734-138	Sequence 138, App
857	35	21.1	705	9	US-10-175-743-598	Sequence 58, App	930	34.5	20.8	489	9	US-09-997-653-138	Sequence 138, App
858	35	21.1	705	9	US-10-176-488-598	Sequence 58, App	931	34.5	20.8	489	9	US-10-174-590-116	Sequence 116, App
859	35	21.1	705	9	US-10-176-492-598	Sequence 58, App	932	34.5	20.8	489	9	US-10-176-758-116	Sequence 116, App
860	35	21.1	705	9	US-10-176-747-598	Sequence 58, App	933	34.5	20.8	489	9	US-10-175-737-116	Sequence 116, App
861	35	21.1	705	9	US-10-176-750-598	Sequence 58, App	934	34.5	20.8	489	9	US-09-993-667-138	Sequence 138, App
862	35	21.1	705	9	US-10-176-985-598	Sequence 58, App	935	34.5	20.8	489	9	US-10-173-706-116	Sequence 116, App
863	35	21.1	705	9	US-10-176-987-598	Sequence 58, App	936	34.5	20.8	489	9	US-10-175-738-116	Sequence 116, App
864	35	21.1	705	9	US-10-176-991-598	Sequence 58, App	937	34.5	20.8	489	9	US-10-175-752-116	Sequence 116, App
865	35	21.1	705	9	US-10-176-992-598	Sequence 58, App	938	34.5	20.8	489	9	US-10-176-482-116	Sequence 116, App
866	35	21.1	705	9	US-10-176-993-598	Sequence 58, App	939	34.5	20.8	489	9	US-10-176-482-116	Sequence 116, App
867	35	21.1	705	9	US-10-184-658-598	Sequence 58, App	940	34.5	20.8	489	9	US-10-180-552-116	Sequence 116, App
868	35	21.1	705	12	US-10-006-867-162	Sequence 162, App	941	34.5	20.8	489	9	US-10-180-557-116	Sequence 116, App
869	35	21.1	705	12	US-10-052-586-598	Sequence 58, App	942	34.5	20.8	489	9	US-09-990-562-138	Sequence 138, App
870	35	21.1	718	10	US-09-529-063-16	Sequence 16, App1	943	34.5	20.8	489	9	US-09-997-428-138	Sequence 138, App
871	35	21.1	720	10	US-09-823-635-2	Sequence 13, App1	944	34.5	20.8	489	9	US-09-997-428-138	Sequence 138, App
872	35	21.1	720	10	US-09-839-408A-13	Sequence 30, App1	945	34.5	20.8	489	9	US-10-173-700-116	Sequence 116, App
873	35	21.1	720	10	US-09-839-408A-30	Sequence 21, App1	946	34.5	20.8	489	9	US-10-173-700-116	Sequence 116, App
874	35	21.1	726	10	US-09-839-408A-21	Sequence 13994, A	947	34.5	20.8	489	9	US-10-174-572-116	Sequence 116, App
875	35	21.1	760	10	US-09-815-242-13994	Sequence 13948, A	948	34.5	20.8	489	9	US-10-174-579-116	Sequence 116, App
876	35	21.1	823	10	US-09-815-242-13348	Sequence 6, App11	949	34.5	20.8	489	9	US-10-174-582-116	Sequence 116, App
877	35	21.1	832	10	US-09-908-322-6	Sequence 2, App11	950	34.5	20.8	489	9	US-10-174-588-116	Sequence 116, App
878	35	21.1	850	10	US-09-985-675-2	Sequence 11310, A	951	34.5	20.8	489	9	US-10-175-739-116	Sequence 116, App
879	35	21.1	856	10	US-09-815-242-11310	Sequence 11310, A	952	34.5	20.8	489	9	US-10-175-740-116	Sequence 116, App
880	35	21.1	856	10	US-09-815-242-1189	Sequence 6, App11	953	34.5	20.8	489	9	US-10-175-740-116	Sequence 116, App
881	35	21.1	874	9	US-09-158-122-6	Sequence 1, App11	954	34.5	20.8	489	9	US-10-176-488-116	Sequence 116, App
882	35	21.1	874	10	US-09-985-675-1	Sequence 10, App1	955	34.5	20.8	489	9	US-10-176-488-116	Sequence 116, App
883	35	21.1	880	10	US-09-923-490-10	Sequence 2, App11	956	34.5	20.8	489	9	US-10-176-492-116	Sequence 116, App
884	35	21.1	948	10	US-09-897-699-2	Sequence 80, App1	957	34.5	20.8	489	9	US-10-176-747-116	Sequence 116, App
885	35	21.1	957	10	US-09-801-574-80	Sequence 340, App	958	34.5	20.8	489	9	US-10-176-750-116	Sequence 116, App
886	35	21.1	1021	9	US-09-895-913A-340	Sequence 5941, App	959	34.5	20.8	489	9	US-10-176-985-116	Sequence 116, App
887	35	21.1	1045	9	US-09-728-626-5941	Sequence 4, App11	960	34.5	20.8	489	9	US-10-176-987-116	Sequence 116, App
888	35	21.1	1050	9	US-10-062-458-4	Sequence 2, App11	961	34.5	20.8	489	9	US-10-176-991-116	Sequence 116, App
889	35	21.1	1382	10	US-09-729-653-2	Sequence 12, App1	962	34.5	20.8	489	9	US-10-176-992-116	Sequence 116, App
890	35	21.1	2710	9	US-10-153-273-12	Sequence 237, App	963	34.5	20.8	489	9	US-10-176-993-116	Sequence 116, App
891	35	21.1	3503	9	US-10-108-605-237	Sequence 1068, App	964	34.5	20.8	489	9	US-10-184-658-116	Sequence 116, App
892	35	21.1	5179	10	US-10-025-380-1068	Sequence 1068, App	965	34.5	20.8	489	10	US-09-988-722-138	Sequence 138, App
893	35	21.1	5179	10	US-09-822-217-1068	Sequence 1068, App	966	34.5	20.8	489	10	US-09-988-723-138	Sequence 138, App
894	35	21.1	5179	10	US-09-833-263-1068	Sequence 707, App	967	34.5	20.8	489	10	US-09-989-727-138	Sequence 138, App
895	35	20.8	150	9	US-10-012-896-707	Sequence 707, App	968	34.5	20.8	489	10	US-09-989-727-138	Sequence 138, App


```
969 34.5 20.8 489 10 US-09-989-731-138 Sequence 138, App
970 34.5 20.8 489 10 US-09-989-732-138 Sequence 138, App
971 34.5 20.8 489 10 US-09-981-073-138 Sequence 138, App
972 34.5 20.8 489 10 US-09-990-442-138 Sequence 138, App
973 34.5 20.8 489 10 US-09-991-163-138 Sequence 138, App
974 34.5 20.8 489 10 US-09-993-604-138 Sequence 138, App
975 34.5 20.8 489 10 US-09-990-456-138 Sequence 138, App
976 34.5 20.8 489 10 US-09-989-721-138 Sequence 138, App
977 34.5 20.8 489 12 US-10-052-586-116 Sequence 116, App
978 34.5 20.8 513 10 US-09-038-079-15 Sequence 15, App
979 34.5 20.8 523 9 US-10-113-790-1 Sequence 1, App
980 34.5 20.8 530 10 US-09-815-242-5570 Sequence 5570, App
981 34.5 20.8 530 10 US-09-815-242-12280 Sequence 12280, App
982 34.5 20.8 530 10 US-09-861-451A-18 Sequence 18, App
983 34.5 20.8 575 10 US-09-815-242-11275 Sequence 11275, App
984 34.5 20.8 581 9 US-09-738-626-6369 Sequence 6269, App
985 34.5 20.8 584 10 US-09-945-825-2 Sequence 2, App
986 34.5 20.8 626 10 US-09-771-161A-201 Sequence 201, App
987 34.5 20.8 626 10 US-09-771-161A-202 Sequence 202, App
988 34.5 20.8 626 12 US-10-000-864-6 Sequence 6, App
989 34.5 20.8 626 12 US-10-000-864-12 Sequence 12, App
990 34.5 20.8 657 10 US-09-815-242-5408 Sequence 5408, App
991 34.5 20.8 667 10 US-09-815-242-12490 Sequence 12490, App
992 34.5 20.8 667 10 US-09-815-242-12811 Sequence 12811, App
993 34.5 20.8 734 9 US-10-170-528-2 Sequence 2, App
994 34.5 20.8 752 9 US-10-003-295-2 Sequence 2, App
995 34.5 20.8 784 10 US-09-905-983-52 Sequence 52, App
996 34.5 20.8 822 9 US-10-003-295-4 Sequence 4, App
997 34.5 20.8 888 10 US-09-223-490-35 Sequence 35, App
998 34.5 20.8 996 9 US-09-738-626-3847 Sequence 3847, App
999 34.5 20.8 1160 9 US-10-028-072-234 Sequence 234, App
1000 34.5 20.8 1160 9 US-10-121-049-234 Sequence 234, App
```

ALIGNMENTS

```
RESULT 1
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8

Query Match 100.0%; Score 166; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 5.7e-17;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
DB 19 SWRSGSFVGLFGQHPDVFYLMPEAMHV 47
```

```
RESULT 2
US-10-007-262-1
; Sequence 1, Application US/10007262
; Patent No. US20020164748A1
; GENERAL INFORMATION:
```

```
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107C1P
; CURRENT APPLICATION NUMBER: US/10/007,262
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 166; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
DB 50 SWRSGSFVGLFGQHPDVFYLMPEAMHV 78
```

```
RESULT 3
US-09-816-825-2
; Sequence 2, Application US/09816825
; Patent No. US20010051370A1
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CON
; CURRENT APPLICATION NUMBER: US/09/816,825
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/045,284
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-825-2
```

```
Query Match 100.0%; Score 166; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 SWRSGSFVGLFGQHPDVFYLMPEAMHV 29
DB 50 SWRSGSFVGLFGQHPDVFYLMPEAMHV 78
```

```
RESULT 4
US-09-927-602-6
; Sequence 6, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
```


SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-6

Query Match 96.4%; Score 160; DB 10; Length 169;
Best Local Similarity 96.6%; Pred. No. 4.1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
Db 16 SWRSGSFVQGLFGQHPDVFLMEPAHV 44

RESULT 5
US-09-927-602-2
Sequence 2, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 395
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-2

Query Match 96.4%; Score 160; DB 10; Length 395;
Best Local Similarity 96.6%; Pred. No. 1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
Db 48 SWRSGSFVQGLFGQHPDVFLMEPAHV 76

RESULT 6
US-09-927-602-3
Sequence 3, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 3
LENGTH: 395
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: synthetic construct
NAME/KEY: VARIANT
LOCATION: (1)...(395)
OTHER INFORMATION: Xaa = any amino acid
US-09-927-602-3

Query Match 96.4%; Score 160; DB 10; Length 395;
Best Local Similarity 96.6%; Pred. No. 1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
Db 48 SWRSGSFVQGLFGQHPDVFLMEPAHV 76

RESULT 7
US-09-927-602-5
Sequence 5, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-927-602-5

Query Match 96.4%; Score 160; DB 10; Length 418;
Best Local Similarity 96.6%; Pred. No. 1.1e-15;
Matches 28; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
Db 72 SWRSGSFVQGLFGQHPDVFLMEPAHV 100

RESULT 8
US-09-927-602-7
Sequence 7, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-7

Query Match 94.6%; Score 157; DB 10; Length 169;
Best Local Similarity 93.1%; Pred. No. 1.1e-15;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSGSFVQGLFGQHPDVFLMEPAHV 29
Db 16 SWRSGSFVQGLFGQHPDVFLMEPAHV 44

RESULT 9

```
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LT 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-4

Query Match          94.6%; Score 157; DB 10; Length 390;
Best Local Similarity 93.1%; Pred. No. 2,7e-15;
Matches 27; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 SWRSSSFVGLFGQHPDYFLMEPAWHV 29
Db 49 SWRSSSFVGLFGQHPDYFLMEPAWHV 77

RESULT 10
US-09-927-602-9
; Sequence 9, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LT 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-9

Query Match          74.7%; Score 124; DB 10; Length 169;
Best Local Similarity 69.0%; Pred. No. 6,7e-11;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSSSFVGLFGQHPDYFLMEPAWHV 29
Db 17 TWRSSSFVGLFGQHPDYFLMEPAWHV 45

RESULT 11
US-10-212-933-2
; Sequence 2, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
```

```
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1.001AUS
; CURRENT APPLICATION NUMBER: US/10/212,933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 483
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-212-933-2

Query Match          74.7%; Score 124; DB 9; Length 483;
Best Local Similarity 69.0%; Pred. No. 2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSSSFVGLFGQHPDYFLMEPAWHV 29
Db 125 TWRSSSFVGLFGQHPDYFLMEPAWHV 153

RESULT 12
US-10-212-933-4
; Sequence 4, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA1.001AUS
; CURRENT APPLICATION NUMBER: US/10/212,933
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; PRIOR FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-933-4

Query Match          74.7%; Score 124; DB 9; Length 484;
Best Local Similarity 69.0%; Pred. No. 2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 SWRSSSFVGLFGQHPDYFLMEPAWHV 29
Db 126 TWRSSSFVGLFGQHPDYFLMEPAWHV 154

RESULT 13
US-09-833-790-255
; Sequence 255, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
```

APPLICANT: Lodges, Michael J.
APPLICANT: Wang, Tonglong
APPLICANT: Secrist, Heather
APPLICANT: Mohamath, Radooh
APPLICANT: Indrias, Carol Y.
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.512
CURRENT APPLICATION NUMBER: US/09/833,790
CURRENT FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 440
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 255
LENGTH: 531
TYPE: PRT
ORGANISM: Homo sapiens
US-09-833-790-255

Query Match 74.7%; Score 124; DB 10; Length 531;
Best Local Similarity 69.0%; Pred. No. 2.2e-10;
Matches 20; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

OY 1 SWRSGSFVGOLFQHPDVFYLMPEPAHV 29
DB 173 TWRSGSFGELEFNQNPVFYLPVHV 201

RESULT 14
US-09-927-602-10
Sequence 10, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
TITLE OF INVENTION: Dystrophy
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 10
LENGTH: 179
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-10

Query Match 69.9%; Score 116; DB 10; Length 179;
Best Local Similarity 81.5%; Pred. No. 1e-09;
Matches 22; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 RSGSFFVGOLFQHPDVFYLMPEPAHV 29
DB 26 RSGSFFVGOLFQHPDVFYLMPEPAHV 52

RESULT 15
US-09-927-602-11
Sequence 11, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
TITLE OF INVENTION: Dystrophy
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11

NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 11
LENGTH: 174
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-11

Query Match 60.8%; Score 101; DB 10; Length 174;
Best Local Similarity 59.3%; Pred. No. 1.5e-07;
Matches 16; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

OY 3 RSGSFFVGOLFQHPDVFYLMPEPAHV 29
DB 22 RTGSSFFGEFFNQNPVFYLPVHV 48

RESULT 16
US-09-738-626-5973
Sequence 5973, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHITO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 5973
LENGTH: 344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5973

Query Match 30.1%; Score 50; DB 9; Length 344;
Best Local Similarity 47.8%; Pred. No. 7.1;
Matches 11; Conservative 2; Mismatches 8; Indels 2; Gaps 1;

OY 8 FVGOLFQHPD--VFYLMPEPAHV 28
DB 187 FYEELKDHDPDAPVYLTGAWN 209

RESULT 17
US-09-888-623-2
Sequence 2, Application US/09888623
Patent No. US20020042115A1
GENERAL INFORMATION:
APPLICANT: Xiao, Yonglong
TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
TITLE OF INVENTION: Synthase Thioesterase-Like Enzyme
FILE REFERENCE: 04974.00510
CURRENT APPLICATION NUMBER: US/09/888,623
CURRENT FILING DATE: 2001-06-26
PRIOR APPLICATION NUMBER: US 60/214,012
PRIOR FILING DATE: 2000-06-26
PRIOR APPLICATION NUMBER: US 60/255,148

```
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-2

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 265;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 219 AMKDVTSGNKITYQLPGH---FYLLDPA 244

RESULT 18
US-09-888-623-14
; Sequence 14, Application US/09888623
; Patent No. US20020042115A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
; FILE REFERENCE: 04974.00510
; CURRENT APPLICATION NUMBER: US/09/888,623
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,012
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/255,148
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-14

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 265;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 219 AMKDVTSGNKITYQLPGH---FYLLDPA 244

RESULT 19
US-09-888-623-16
; Sequence 16, Application US/09888623
; Patent No. US20020042115A1
; GENERAL INFORMATION:
; APPLICANT: Xiao, Yonghong
; TITLE OF INVENTION: Regulation of Human S-Acyl Fatty Acid
; FILE REFERENCE: 04974.00510
; CURRENT APPLICATION NUMBER: US/09/888,623
; PRIOR FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: US 60/214,012
; PRIOR FILING DATE: 2000-06-26
; PRIOR APPLICATION NUMBER: US 60/255,148
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: PCT 01/06936
; PRIOR FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 16
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-623-16

Query Match
Best Local Similarity 41.4%; Score 49; DB 10; Length 318;
Matches 12; Conservative 7; Mismatches 4; Indels 6; Gaps 2;

QY 1 SWR---SGSFFVGLFGQHPDYFLMEPA 26
Db 272 AMKDVTSGNKITYQLPGH---FYLLDPA 297

RESULT 20
US-09-815-242-5057
; Sequence 5057, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5057
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5057

Query Match
Best Local Similarity 33.3%; Score 48; DB 10; Length 460;
Matches 8; Conservative 6; Mismatches 10; Indels 0; Gaps 0;

QY 5 GSSFFVGLFGQHPDYFLMEPAWH 28
Db 169 GSKFHEHMGMLPFAHIDEPTWY 192

RESULT 21
US-09-888-615-68
; Sequence 68, Application US/09888615
; Patent No. US20020064856A1
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEL, SEAN
; APPLICANT: CHARVETZAK, GLEN
```



```

; ORGANISM: Corynebacterium glutamicum
US-09-738-626-3745

Query Match          27.7%; Score 46; DB 9; Length 302;
Best Local Similarity 43.8%; Pred. No. 23;
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 14 GQHPDVFYLMPEAMV 29
      |:::|::|::|::|
Db 179 GEHEKMSIYEPPMHV 194

RESULT 25
US-09-272-809-6
; Sequence 6, Application US/09272809
; Patent No. US20020022239A1
; GENERAL INFORMATION:
; APPLICANT: Lagarias, John C.
; TITLE OF INVENTION: Phytofluors as fluorescent labels
; FILE REFERENCE: 2500.118USO
; CURRENT APPLICATION NUMBER: US/09/272,809
; CURRENT FILING DATE: 1998-03-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism:cpb6 locus
; OTHER INFORMATION: SLR12112 (ETRI homolog, PAS domain) an 844 aa
US-09-272-809-6

Query Match          27.4%; Score 45.5; DB 10; Length 844;
Best Local Similarity 30.8%; Pred. No. 82;
Matches 8; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

QY 3 RSGSSVVGQLFGQHPDVFILMEP-AM 27
      |::|::|::|::|::|::|
Db 321 RESQYFIETKLANYSFQILYLDFIAW 346

RESULT 26
US-09-864-761-41772
; Sequence 41772, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anncmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41772
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC015971.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.2
; OTHER INFORMATION: EST_HUMAN HIT: BF218030.1, EVALUE 1.00e-02
; OTHER INFORMATION: SWISSPROT HIT: Q05793, EVALUE 6.40e+00
US-09-864-761-41772

Query Match          27.1%; Score 45; DB 10; Length 99;
Best Local Similarity 53.3%; Pred. No. 10;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 2 WRSGSSVVGQLFGQH 16
      |::|::|::|::|
Db 73 WRNSSTLGYLFESH 87

RESULT 27
US-09-925-299-834
; Sequence 834, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 834
; LENGTH: 239
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-925-299-834

Query Match          27.1%; Score 45; DB 10; Length 239;
Best Local Similarity 33.3%; Pred. No. 26;
```

	Matches	9;	Conservative	3;	Mismatches	15;	Indels	0;	Gaps	0;
QY	1	SWRSGSFVQGLFGQHPDVEYLMEPAW	27							
				:	:	:				
Db	65	SWRPSVEFPNGNLYRGEGLIYGGLEEW	91							

```

RESULT 28
US-09-738-626-6189
: Sequence 6189, Application US/09738626
: Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIKO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKTO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738, 626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 6189
LENGTH: 376
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6189

```

	Query Match	27.1%	Score 45;	DB 9;	Length 376;
	Best Local Similarity	47.8%;	Pred.No. 41;		
Matches	11; Conservative	1;	Mismatches	11;	Gaps 0;
OY	5 GSFVYGQLFGQHPDVFTIMEPAW	27			
Db	59 GSSOVGMIMGYHRAGFSGMALAW	81			

```

RESULT 29
US-09-735-705-170
: Sequence No. 170, Application US/09735705
: Patent No. US20020052329A1
:
: GENERAL INFORMATION:
:
: APPLICANT: Wang, Tongtong
: APPLICANT: Fan, Liqun
: APPLICANT: Kalos, Michael D.
: APPLICANT: Bangur, Chaitanya S.
: APPLICANT: Hosken, Nancy
: APPLICANT: Fanger, Gary R.
: APPLICANT: Li, Samuel X.
: APPLICANT: Wang, Aijun
: APPLICANT: Skeiky, Yashir A.W.
: APPLICANT: Henderson, Robert A.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Fanger, Neil
:
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
:
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
:
: FILE REFERENCE: 210121.455C14
:
: CURRENT APPLICATION NUMBER: US/09/735,705
:
: CURRENT FILING DATE: 2000-12-12
:
: NUMBER OF SEQ ID NOS: 419
:
: SOFTWARE: FASTSEQ for Windows Version 3.0

```

```

; SEQ ID NO 170
; LENGTH: 791
; TYPE: prt
; ORGANISM: Homo sapien
US-09-735-705-170

```

Query Match	27.1%;	Score 45;	DB 10;	Length 791;
Best Local Similarity	57.9%;	Pred. No. 91;		
Matches 11;	Conservative 1;	Mismatches 5;	Indels 2;	Gaps 1;

```

: RESULT 30
: US-09-850-716A-170
: Sequence 170, Application US/09850716A
: Patent No. US2009015139A1
: GENERAL INFORMATION:
: APPLICANT: Kalos, Michael D.
: APPLICANT: McNeill, Patricia D.
: APPLICANT: Retter, Marc W.
: TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
: TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
: FILE REFERENCE: 210121.455C15
: CURRENT APPLICATION NUMBER: US/09/850,716A
: CURRENT FILING DATE: 2001-05-07
: NUMBER OF SEQ ID NOS: 440
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 170
: LENGTH: 791
: TYPE: PRT
: ORGANISM: Homo sapien
: US-09-850-716A-170

```

	Query Match	27.1%	Score 45:	DB 10;	Length 791;
	Best Local Similarity	57.9%;	Pred. No.	91;	
	Matches 11;	Conservative	1;	Mismatches	5; Indels 2; Gaps 1;
QY	4 SSSS--VGQLFGQHDPVF	20			
Db	746 SGGSSVLLGVPAGPHDPVF	764			

```

RESULT 31
US-09-897-778-170
Sequence 170 Application US/09897778
Patent No. US20020147143A1
GENERAL INFORMATION:
APPLICANT: Wang, Tonglong
APPLICANT: Marnierakis, Margarita
APPLICANT: Fanger, Gary R.
APPLICANT: Vedrick, Thomas S.
APPLICANT: Carter, Darrick
APPLICANT: Watanabe, Yoshihiro
APPLICANT: Henderson, Robert A.
APPLICANT: Peckham, David W.
APPLICANT: Fanger, Neil
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
FILE REFERENCE: 210121.455C16
CURRENT APPLICATION NUMBER: US/09/897,778
CURRENT FILING DATE: 2001-06-28
NUMBER OF SEQ ID NOS: 467
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 170
LENGTH: 791
TYPE: PRT
ORGANISM: Homo sapiens
US-09-897-778-170
Query Match 27.1%; Score 45; DB 10; Length 791

```



```
; Sequence 357, Application US/09850716A
; Patent No. US2002011519A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Retter, Marc W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850,716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 357
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-850-716A-357
```

```
Query Match          27.1%  Score 45; DB 10; Length 920;
Best Local Similarity 57.9%  Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGSSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      723  SGGSFVLGVAPGPHDPVF 741
```

```
RESULT 37
US-09-897-778-357
; Sequence 357, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darlick
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897,778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
; LENGTH: 920
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-778-357
```

```
Query Match          27.1%  Score 45; DB 10; Length 920;
Best Local Similarity 57.9%  Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGSSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      723  SGGSFVLGVAPGPHDPVF 741
```

```
RESULT 38
US-09-974-298-71
; Sequence 71, Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Huel-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
```

```
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 71
; LENGTH: 942
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 2733282CD1
US-09-974-298-71
```

```
Query Match          27.1%  Score 45; DB 9; Length 942;
Best Local Similarity 57.9%  Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGSSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      745  SGGSFVLGVAPGPHDPVF 763
```

```
RESULT 39
US-09-919-172-87
; Sequence 87, Application US/09919172
; Patent No. US20020119463A1
; GENERAL INFORMATION:
; APPLICANT: Faris, Mary
; APPLICANT: Turner, Christopher M.
; TITLE OF INVENTION: PROSTATE CANCER MARKERS
; FILE REFERENCE: PA-0036 US
; CURRENT APPLICATION NUMBER: US/09/919,172
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/222,469
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PERL Program
; SEQ ID NO 87
; LENGTH: 942
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020119463A1 2733282CD1
US-09-919-172-87
```

```
Query Match          27.1%  Score 45; DB 10; Length 942;
Best Local Similarity 57.9%  Pred. No. 1.1e+02;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;
```

```
QY      4  SGSSF--VGQLFGQHPDVF 20
      || || :| :| |||||
Db      745  SGGSFVLGVAPGPHDPVF 763
```

```
RESULT 40
US-09-735-705-161
; Sequence 161, Application US/09735705
; Patent No. US20020052329A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy
; APPLICANT: Fanger, Gary R.
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Henderson, Robert A.
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Fanger, Neil
```

```
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C14
; CURRENT APPLICATION NUMBER: US/09/735.705
; CURRENT FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 419
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRN
; ORGANISM: Homo sapien
US-09-735-705-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGQHPDYF 20
DB 746 SSSSFVGLGVPAGHPDYF 764

RESULT 41
US-09-850-716A-161
; Sequence 161, Application US/09850716A
; Patent No. US20020115139A1
; GENERAL INFORMATION:
; APPLICANT: Kalos, Michael D.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C15
; CURRENT APPLICATION NUMBER: US/09/850.716A
; CURRENT FILING DATE: 2001-05-07
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRN
; ORGANISM: Homo sapien
US-09-850-716A-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGQHPDYF 20
DB 746 SSSSFVGLGVPAGHPDYF 764

RESULT 42
US-09-897-778-161
; Sequence 161, Application US/09897778
; Patent No. US20020147143A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tonglong
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedyick, Thomas S.
; APPLICANT: Carter, Darick
; APPLICANT: Matanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Peckham, David W.
; APPLICANT: Fanger, Neil
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.455C16
; CURRENT APPLICATION NUMBER: US/09/897.778
; CURRENT FILING DATE: 2001-06-28
; NUMBER OF SEQ ID NOS: 467
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 161
; LENGTH: 943
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-897-778-161

Query Match
Best Local Similarity 27.1%; Score 45; DB 10; Length 943;
Matches 11; Conservative 1; Mismatches 5; Indels 2; Gaps 1;

QY 4 SSSSF--VGQLFGQHPDYF 20
DB 746 SSSSFVGLGVPAGHPDYF 764

RESULT 43
US-09-815-242-11294
; Sequence 11294, Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815.242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11294
; LENGTH: 812
; TYPE: PRN
; ORGANISM: Helicobacter pylori
US-09-815-242-11294

Query Match
Best Local Similarity 26.5%; Score 44; DB 10; Length 812;
Matches 9; Conservative 5; Mismatches 9; Indels 2; Gaps 1;

QY 1 SRRSSSFVQQLFGQHPDYFLMEP 25
DB 220 AWMGLGVNNGT--NPDEFVFKP 242

RESULT 44
US-09-738-626-5852
; Sequence 5852, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
```

APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAKO
APPLICANT: SENO, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OKAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO 5852
LENGTH: 1054
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-5852

Query Match 26.2%; Score 43.5; DB 9; Length 1054;
Best Local Similarity 31.0%; Pred. No. 2e+02;
Matches 9; Conservative 3; Mismatches 6; Indels 11; Gaps 1;

QY 1 SWRSGSFVGLFGQHPDVFYLMPEANHV 29
Db 422 SWRSGEPLI-----YMALPSWFV 439

RESULT 45
US-09-925-302-729
Sequence 729, Application US/09925502
Patent No. US20020044941A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
FILE REFERENCE: PAL06
CURRENT APPLICATION NUMBER: US/09/925,302
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: PCT/US00/05918
PRIOR FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/124,270
PRIOR FILING DATE: 1999-03-12
NUMBER OF SEQ ID NOS: 896
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 729
LENGTH: 466
TYPE: PRT
ORGANISM: Homo sapiens
US-09-925-302-729

Query Match 25.9%; Score 43; DB 10; Length 466;
Best Local Similarity 37.0%; Pred. No. 1e+02;
Matches 10; Conservative 6; Mismatches 7; Indels 4; Gaps 2;

QY 4 SGSSFVGLFGQHPDVFYLMPEANHV 29
Db 398 AASTFNG---FQRPITYVMSGPAWQL 421

RESULT 46
US-09-815-242-5480
Sequence 5480, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel

APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5480
LENGTH: 1160
TYPE: PRT
ORGANISM: Staphylococcus aureus
US-09-815-242-5480

Query Match 25.9%; Score 43; DB 10; Length 1160;
Best Local Similarity 42.9%; Pred. No. 2.7e+02;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 4 SGSSFVGLFGQHPDVFYLMPEANHV 24
Db 332 SGNGFVGSGFQKXDFEFLIE 352

RESULT 47
US-09-815-242-12151
Sequence 12151, Application US/09815242
Patent No. US20020061569A1
GENERAL INFORMATION:
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Karl L.
APPLICANT: Zyskind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.
APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
TITLE OF INVENTION: Prokaryotes
FILE REFERENCE: ELITRA.011A
CURRENT APPLICATION NUMBER: US/09/815,242
CURRENT FILING DATE: 2001-03-21
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16
NUMBER OF SEQ ID NOS: 14110

```
SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12151
; LENGTH: 1168
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12151
```

```
Query Match
Best Local Similarity 42.9%; DB 10; Length 1168;
Matches 9; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
```

```
OY 4 SGGSFVGLFGQHPDYFLMEP 24
||:||||:|:|:|:|
DB 332 SGGSFVGLFGQHPDYFLMEP 352
```

```
RESULT 48
US-09-764-870-510
; Sequence 510, Application US/09764870
; Patent No. US2002042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT214
; CURRENT APPLICATION NUMBER: US/09764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 510
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-510
```

```
Query Match
Best Local Similarity 47.8%; DB 10; Length 80;
Matches 11; Conservative 0; Mismatches 9; Indels 3; Gaps 1;
```

```
OY 3 RGGSPFVGLFGQHPDYFLMEP 25
|||||:|:|:|:|
DB 31 RGGSPFVGLFGQHPDYFLMEP 50
```

```
RESULT 49
US-10-001-835-184
; Sequence 184, Application US/10001835
; Patent No. US20020160387A1
; GENERAL INFORMATION:
; APPLICANT: Salceda, Susana
; APPLICANT: Macina, Roberto
; APPLICANT: Recipon, Herve
; APPLICANT: Caferkey, Robert
; APPLICANT: Sun, Yongming
; APPLICANT: Liu, Chenghua
; TITLE OF INVENTION: Compositions and Methods Relating to Ovary Specific Genes and Prc
; FILE REFERENCE: DEX-0277
; CURRENT APPLICATION NUMBER: US/10/001,835
; CURRENT FILING DATE: 2001-11-20
; PRIOR APPLICATION NUMBER: 60/249,997
; PRIOR FILING DATE: 2000-11-20
; NUMBER OF SEQ ID NOS: 228
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 184
; LENGTH: 94
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-001-835-184
```

```
Query Match
Best Local Similarity 25.6%; Score 42.5; DB 9; Length 94;
Matches 12; Conservative 2; Mismatches 9; Indels 13; Gaps 2;
```

```
OY 1 SWRGSSFFVG-----OLFQHPDYFLMEPAW 27
|||:|:|:|:|:|:|:|:|:|
DB 60 SWRGSSFFVG-----OLFQHPDYFLMEPAW 91
```

```
RESULT 50
US-09-992-598-281
; Sequence 281, Application US/09992598
; Patent No. US20020160384A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi J.
; APPLICANT: Baker, Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Fong, Sherman
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2730P1C20
; CURRENT APPLICATION NUMBER: US/09/992,598
; CURRENT FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: 60/049787
; PRIOR FILING DATE: 1997-06-16
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/075945
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/087106
; PRIOR FILING DATE: 1998-05-28
; PRIOR APPLICATION NUMBER: 60/087607
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087609
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087759
; PRIOR FILING DATE: 1998-06-02
; PRIOR APPLICATION NUMBER: 60/087827
; PRIOR FILING DATE: 1998-06-03
; PRIOR APPLICATION NUMBER: 60/088021
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088025
; PRIOR FILING DATE: 1998-06-04
; PRIOR APPLICATION NUMBER: 60/088026
; PRIOR FILING DATE: 1998-06-04
```

PRIOR APPLICATION NUMBER: 60/088028
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088029
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088030
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088033
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088326
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088167
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088202
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088212
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088217
PRIOR FILING DATE: 1998-06-05
PRIOR APPLICATION NUMBER: 60/088555
PRIOR FILING DATE: 1998-06-09
PRIOR APPLICATION NUMBER: 60/088734
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088738
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088742
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088824
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088826
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088861
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/088876
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089105
PRIOR FILING DATE: 1998-06-12
PRIOR APPLICATION NUMBER: 60/089440
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089512
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089514
PRIOR FILING DATE: 1998-06-16
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089598
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089600
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089653
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089801
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089908
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089948
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/089952
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090246
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090252

PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090254
PRIOR FILING DATE: 1998-06-22
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090355
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090431
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090435
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090444
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090535
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090540
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090542
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090676
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090678
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090690
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090694
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090696
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090862
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091478
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091544
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091626
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091633
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091978
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/092182
PRIOR FILING DATE: 1998-07-09

Query Match 25.6% Score 42.5; DB 9; Length 229;
Best Local Similarity 37.0%; Pred. No. 56;
Matches 10; Conservative 2; Mismatches 8; Indels 7; Gaps 1;

QY 2 WSGSSVWGOLF-----GQHPDVY 21
DB 118 WSGSDEAPGLFFVDAERVPCRHDDVFF 144

Search completed: February 20, 2003, 13:40:48
Job time : 29.2571 secs

is Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:26:11 ; Search time 10.6286 Seconds
(Without alignments)
75.222 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table:

BLOSUM62
Gap 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :
1: A_Geneseq.101002.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	169	23	ABB81561 Human N-acetylgluc
2	31	100.0	171	23	ABB81560 Human high endoth
3	31	100.0	179	23	ABB81562 Human keratan sulf
4	31	100.0	183	22	ABB68582 Drosophila melanog
5	31	100.0	380	23	AA011274 Human L-selectin s
6	31	100.0	386	20	AA139918 Human glycosyl sul
7	31	100.0	386	21	AA179219 Human transferrase
8	31	100.0	386	22	AA193309 Human polyepitide
9	31	100.0	395	22	AA172638 Mouse glycosyl sul
10	31	100.0	395	23	AA011275 Murine intestinal-

11	31	100.0	411	19	AAW61100 Keratan sulphate 6
12	31	100.0	418	23	ABB81557 Mouse intestinal N
13	31	100.0	483	20	AA131656 Mouse N-acetylgluc
14	31	100.0	484	20	AA131657 Human N-acetylgluc
15	31	100.0	530	22	AA193367 Human protein sequ
16	31	100.0	531	22	AA069414 Lung small cell ca
17	31	100.0	649	22	AA067037 Drosophila melanog
18	30	96.8	363	22	ABB64512 Drosophila melanog
19	30	96.8	562	23	ABB09452 Glucose-3-dehydrog
20	29	93.5	405	23	ABB92113 Herbicidally activ
21	28	90.3	139	23	ABB32935 Human isomerase-11
22	28	90.3	174	23	ABB81563 Human Chondroitin-
23	28	90.3	253	22	ABB55144 Lactococcus lactis
24	28	90.3	306	17	Escherichia coli P
25	28	90.3	306	17	Capsular polysacch
26	28	90.3	306	20	AA133895 Amino acid sequenc
27	28	90.3	333	23	ABB91753 Herbicidally activ
28	28	90.3	349	21	AA14207 Arabidopsis thalia
29	28	90.3	349	21	AA14207 Arabidopsis thalia
30	28	90.3	354	21	AA14206 Arabidopsis thalia
31	28	90.3	354	21	AA14206 Arabidopsis thalia
32	28	90.3	354	21	AA14206 Arabidopsis thalia
33	28	90.3	359	21	AA14205 Arabidopsis thalia
34	28	90.3	359	21	AA14205 Arabidopsis thalia
35	28	90.3	359	21	AA14205 Arabidopsis thalia
36	28	90.3	388	20	AA13919 Hydroxyjasmonic ac
37	28	90.3	458	18	AA13919 Mouse glycosyl sul
38	28	90.3	458	18	AA13919 Chick chondroitin
39	28	90.3	479	19	AA13919 Putative P. abyssal
40	28	90.3	567	15	AA13919 Glycosaminoglycan
41	28	90.3	612	23	ABB92573 Mammalian MEK kina
42	28	90.3	626	16	AA177546 MEK3 protein. Mu
43	28	90.3	626	19	AA177546 MEK3 protein. Mu
44	28	90.3	626	20	AA177546 MEK3 protein. Mu
45	28	90.3	626	20	AA177546 MEK3 protein. Mu
46	28	90.3	626	20	AA177546 MEK3 protein. Mu
47	28	90.3	626	20	AA177546 MEK3 protein. Mu
48	28	90.3	626	20	AA177546 MEK3 protein. Mu
49	28	90.3	626	20	AA177546 MEK3 protein. Mu
50	28	90.3	626	20	AA177546 MEK3 protein. Mu
51	28	90.3	626	20	AA177546 MEK3 protein. Mu
52	28	90.3	626	20	AA177546 MEK3 protein. Mu
53	28	90.3	626	20	AA177546 MEK3 protein. Mu
54	28	90.3	626	20	AA177546 MEK3 protein. Mu
55	28	90.3	626	20	AA177546 MEK3 protein. Mu
56	28	90.3	626	20	AA177546 MEK3 protein. Mu
57	28	90.3	626	20	AA177546 MEK3 protein. Mu
58	28	90.3	626	20	AA177546 MEK3 protein. Mu
59	28	90.3	626	20	AA177546 MEK3 protein. Mu
60	28	90.3	626	20	AA177546 MEK3 protein. Mu
61	28	90.3	626	20	AA177546 MEK3 protein. Mu
62	28	90.3	626	20	AA177546 MEK3 protein. Mu
63	28	90.3	626	20	AA177546 MEK3 protein. Mu
64	28	90.3	626	20	AA177546 MEK3 protein. Mu
65	28	90.3	626	20	AA177546 MEK3 protein. Mu
66	28	90.3	626	20	AA177546 MEK3 protein. Mu
67	28	90.3	626	20	AA177546 MEK3 protein. Mu
68	28	90.3	626	20	AA177546 MEK3 protein. Mu
69	28	90.3	626	20	AA177546 MEK3 protein. Mu
70	28	90.3	626	20	AA177546 MEK3 protein. Mu
71	28	90.3	626	20	AA177546 MEK3 protein. Mu
72	28	90.3	626	20	AA177546 MEK3 protein. Mu
73	28	90.3	626	20	AA177546 MEK3 protein. Mu
74	28	90.3	626	20	AA177546 MEK3 protein. Mu
75	28	90.3	626	20	AA177546 MEK3 protein. Mu
76	28	90.3	626	20	AA177546 MEK3 protein. Mu
77	28	90.3	626	20	AA177546 MEK3 protein. Mu
78	28	90.3	626	20	AA177546 MEK3 protein. Mu
79	28	90.3	626	20	AA177546 MEK3 protein. Mu
80	28	90.3	626	20	AA177546 MEK3 protein. Mu
81	28	90.3	626	20	AA177546 MEK3 protein. Mu
82	28	90.3	626	20	AA177546 MEK3 protein. Mu
83	28	90.3	626	20	AA177546 MEK3 protein. Mu

84	27	87.1	308	18	AAW10669	Human TGF-beta-1ik	157	26	83.9	148	21	AA61860	Arabidopsis thalia
85	27	87.1	308	18	AAW10670	Human TGF-beta-1ik	158	26	83.9	148	23	AAU72540	Arabidopsis cell c
86	27	87.1	308	18	AAW10671	Human TGF-beta-1ik	159	26	83.9	155	22	AAU17354	Novel signal trans
87	27	87.1	308	19	AAW48672	GF-2H protein, Ho	160	26	83.9	170	23	AAU72512	Arabidopsis paria
88	27	87.1	308	21	AAW26122	Human TGF-beta sup	161	26	83.9	249	22	ABW69168	Drosophila melanog
89	27	87.1	308	23	ABW09174	Human NAG-1 protei	162	26	83.9	253	23	ABW61509	Iron uptake ABC tr
90	27	87.1	308	23	AAE13538	Human macrophage 1	163	26	83.9	315	22	AAE10324	Human transporter
91	27	87.1	308	23	AAE13539	Human macrophage 1	164	26	83.9	315	22	AAW94269	Human protein sequ
92	27	87.1	316	21	AAW19726	Human SECX Clone 4	165	26	83.9	315	22	AAW94664	Human protein sequ
93	27	87.1	318	18	AAW10663	Human TGF-beta-1ik	166	26	83.9	332	22	AAU17660	Novel signal trans
94	27	87.1	325	22	ABW57775	Drosophila melanog	167	26	83.9	346	22	ABW62612	Drosophila melanog
95	27	87.1	330	22	AAW72639	Human glycosyl sul	168	26	83.9	407	22	AAW91787	C glutamicum prote
96	27	87.1	330	23	ABW1556	Human intestinal N	169	26	83.9	417	13	AAW27725	Sequence transcrib
97	27	87.1	335	22	AAW72640	Human glycosyl sul	170	26	83.9	423	21	AAW03064	Maize KIN17 orthol
98	27	87.1	335	23	ABW1554	Human corneal N-ac	171	26	83.9	423	21	AAW03065	Propionibacterium
99	27	87.1	335	23	ABW1555	Consensus N-acetyl	172	26	83.9	424	22	AAU47211	Hericicidally activ
100	27	87.1	335	23	AAE15438	Human drug metabol	173	26	83.9	449	23	ABW93640	Novel human diagno
101	27	87.1	336	22	AAW36948	Staphylococcus aur	174	26	83.9	463	18	AAW37445	Novel human diagno
102	27	87.1	336	22	AAU37307	Staphylococcus aur	175	26	83.9	465	18	AAW37438	Rat islet glucokin
103	27	87.1	405	18	AAW20788	H. pylori cytoplas	176	26	83.9	465	18	AAW37439	Rat liver glucokin
104	27	87.1	418	21	AAW41947	Human ORFX ORF1711	177	26	83.9	465	18	AAW37431	Rat liver glucokin
105	27	87.1	426	22	ABW15281	Novel human diagno	178	26	83.9	465	18	AAW37432	Rat liver glucokin
106	27	87.1	431	23	AAU29327	Streptococcus poly	179	26	83.9	542	22	AAW85387	Human protein sequ
107	27	87.1	439	22	AAU34691	E. coli cellular p	180	26	83.9	662	20	AAW34751	C. pneumoniae prot
108	27	87.1	439	22	AAW98252	Escherichia coli p	181	26	83.9	683	22	ABW21469	Novel human diagno
109	27	87.1	476	22	ABW17759	Novel human diagno	182	26	83.9	724	23	AAW22978	Rice MHI protein.
110	27	87.1	488	23	ABW53898	Lactococcus lactis	183	26	83.9	727	22	AAW92265	Human polypeptide
111	27	87.1	499	16	AAW64269	Arabidopsis epsilon-c	184	26	83.9	739	22	ABW67865	Drosophila melanog
112	27	87.1	511	16	AAW41958	Human polypeptide	185	26	83.9	751	22	ABW60916	Drosophila melanog
113	27	87.1	543	23	AAU11262	Novel protein enco	186	26	83.9	753	22	AAW41051	Human polypeptide
114	27	87.1	551	21	AAW52398	Human keratin KERT	187	26	83.9	757	22	ABW58138	Drosophila melanog
115	27	87.1	551	23	AAE20423	Human keratin-2 (K	188	26	83.9	761	22	ABW21470	Novel human diagno
116	27	87.1	562	23	ABW91595	Herbicidally activ	189	26	83.9	838	21	AAW35855	Arabidopsis thalia
117	27	87.1	568	20	AAW35655	Protein involved i	190	26	83.9	843	21	AAW35854	Arabidopsis thalia
118	27	87.1	604	23	AAE18106	Human nucleoside p	191	26	83.9	849	21	AAW35853	Arabidopsis thalia
119	27	87.1	605	22	ABW32651	Escherichia coli p	192	26	83.9	904	23	AAW020516	Protein of APP rel
120	27	87.1	606	21	AAW56855	Semaphorin K1 poly	193	26	83.9	919	18	AAW37440	Hexokinase-glucoki
121	27	87.1	653	22	AAW47250	Human P7. Hom s	194	26	83.9	919	18	AAW37434	Hexokinase-glucoki
122	27	87.1	666	23	ABW47488	Listeria monocytog	195	26	83.9	919	18	AAW37436	Hexokinase-glucoki
123	27	87.1	671	23	ABW41598	Human ovarian anti	196	26	83.9	919	18	AAW37433	Hexokinase-glucoki
124	27	87.1	706	22	ABW65428	Drosophila melanog	197	26	83.9	1029	14	AAW38862	GC-A. Rattus ratt
125	27	87.1	815	21	AAW18728	Human SECX Clone A	198	26	83.9	1171	22	AAW32421	Novel human secret
126	27	87.1	815	22	AAW08678	Human FCTR2 polype	199	26	83.9	1174	11	AAW08257	B. thuringiensis to
127	27	87.1	819	22	AAW78521	Human protein SEO	200	26	83.9	1174	13	AAW25825	Novel toxin expres
128	27	87.1	838	22	AAW79505	Human protein SEO	201	26	83.9	1174	17	AAW89494	B.t. toxin 81A2.
129	27	87.1	842	21	AAW19727	Human SECX Clone 4	202	26	83.9	1174	18	AAW09043	Bacillus thuringie
130	27	87.1	842	22	AAW82300	Human follistatin-	203	26	83.9	1348	22	ABW60761	Drosophila melanog
131	27	87.1	851	22	AAW78324	Human protein SEO	204	26	83.9	1396	20	AAW36871	Protein involved i
132	27	87.1	983	22	AAW70775	Follistatin-relate	205	26	83.9	2464	22	ABW63174	Drosophila melanog
133	27	87.1	983	23	AAE22858	Human Zfsia2 prote	206	26	83.9	2548	20	AAW5761	Human myosin IXa.
134	27	87.1	1119	22	AAW97825	Thermus aquaticus	207	26	80.6	39	21	AAW81628	Cone snail alpha-c
135	27	87.1	1119	23	AAE14716	Chlamydia pneumonia	208	26	80.6	52	21	AAW08204	Arabidopsis thalia
136	27	87.1	1218	20	AAW34680	Human polypeptide	209	26	80.6	55	22	AAW39137	Human colon cancer
137	27	87.1	1266	22	AAW40172	Human polypeptide	210	26	80.6	65	22	AAW39137	Propionibacterium
138	27	87.1	1833	22	ABW71141	Drosophila melanog	211	26	80.6	72	21	AAW08203	Arabidopsis thalia
139	27	87.1	2257	22	ABW61782	Utrrophin. Homo sa	212	26	80.6	88	18	AAW28240	Amino acid sequenc
140	27	87.1	3433	18	AAW22017	Peptide #1359 enco	213	26	80.6	102	21	AAW07415	Arabidopsis thalia
141	26	83.9	45	22	ABW26868	Peptide #1378 enco	214	26	80.6	102	21	AAW32541	Arabidopsis thalia
142	26	83.9	45	22	ABW3872	Protein #1310 enco	215	26	80.6	103	21	AAW49433	Arabidopsis thalia
143	26	83.9	45	22	AAW19311	Human brain expres	216	26	80.6	103	22	AAW72397	Rice proline trans
144	26	83.9	45	22	AAW54637	Human bone marrow	217	26	80.6	104	21	AAW07414	Arabidopsis thalia
145	26	83.9	45	22	AAW67043	Peptide #1377 enco	218	26	80.6	104	21	AAW32540	Arabidopsis thalia
146	26	83.9	45	22	AAW14903	Peptide #1369 enco	219	26	80.6	104	21	AAW41699	Arabidopsis thalia
147	26	83.9	45	22	AAW27332	Peptide #1310 enco	220	26	80.6	105	21	AAW49432	Arabidopsis thalia
148	26	83.9	45	22	AAW02628	Arabidopsis thalia	221	26	80.6	112	21	AAW52616	Helicobacter pylor
149	26	83.9	130	21	AAW06258	Arabidopsis thalia	222	26	80.6	122	21	AAW07413	Arabidopsis thalia
150	26	83.9	130	21	AAW06258	Arabidopsis thalia	223	26	80.6	122	21	AAW32539	Arabidopsis thalia
151	26	83.9	130	21	AAW06258	Arabidopsis thalia	224	26	80.6	122	21	AAW41698	Arabidopsis thalia
152	26	83.9	135	21	AAW06257	Arabidopsis thalia	225	26	80.6	123	21	AAW49431	Arabidopsis thalia
153	26	83.9	135	21	AAW06257	Arabidopsis thalia	226	26	80.6	128	11	AAW06252	Variable region of
154	26	83.9	144	22	AAU18576	Human lung antigen	227	26	80.6	129	18	AAW20536	H. pylori transmem
155	26	83.9	144	22	AAU18543	Novel human respir	228	26	80.6	133	21	AAW41697	Arabidopsis thalia
156	26	83.9	148	21	AAW06256	Arabidopsis thalia	229	26	80.6	139	22	AAW01483	Human polypeptide

230	25	80.6	142	22	ABR65566	Drosophila melanog
231	25	80.6	146	22	AAU04566	Human G-protein co
232	25	80.6	171	22	AAE03594	Wheat tryptophanyl
233	25	80.6	182	20	AAZ56586	Cockroach allergen
234	25	80.6	182	20	AAW88108	A cockroach allergy
235	25	80.6	185	23	AAW62799	Eubacterial DNA po
236	25	80.6	196	18	AAW20596	H. pylori membrane
237	25	80.6	196	18	AAW20596	Human olfactory re
238	25	80.6	243	18	AAW09423	Baiana polyphenol
239	25	80.6	248	14	AAW36280	HML1. Homo sapien
240	25	80.6	248	21	AAW39090	Arabidopsis thalia
241	25	80.6	248	21	AAZ58413	14-3-3 sigma amino
242	25	80.6	249	19	AAW60198	Human mammary epit
243	25	80.6	251	22	ABR68580	Drosophila melanog
244	25	80.6	274	22	AAW62696	Propionibacterium
245	25	80.6	284	22	AAW66362	Putative P. abyssi
246	25	80.6	300	22	AAW60924	Propionibacterium
247	25	80.6	301	21	AAW39089	Arabidopsis thalia
248	25	80.6	305	23	AAW61514	Iron uptake ABC tr
249	25	80.6	313	22	AAW05861	Pseudomonas stutze
250	25	80.6	316	22	ABR66006	Drosophila melanog
251	25	80.6	317	19	AAW60997	Streptococcus pneu
252	25	80.6	321	23	ABR84748	DNA polymerase III
253	25	80.6	322	23	ABP38349	Staphylococcus epi
254	25	80.6	328	22	ABR67466	Drosophila melanog
255	25	80.6	332	22	AAW46741	H. pylori DnaB pro
256	25	80.6	333	23	ABR93585	Herbicidally activ
257	25	80.6	340	23	ABR93581	Herbicidally activ
258	25	80.6	347	22	AAW49723	Hydroxyjaconic ac
259	25	80.6	371	22	AAU35323	Enterococcus faeca
260	25	80.6	373	22	AAW89814	C. glutamicum prote
261	25	80.6	374	22	AAW83001	S. epidermidis ope
262	25	80.6	377	22	AAU27595	Fusion protein 730
263	25	80.6	383	21	AAW09654	Arabidopsis thalia
264	25	80.6	383	21	AAW13572	Arabidopsis thalia
265	25	80.6	383	21	AAW17445	Arabidopsis thalia
266	25	80.6	383	21	AAW36744	Arabidopsis thalia
267	25	80.6	383	21	AAW45472	Arabidopsis thalia
268	25	80.6	383	21	AAW45500	Arabidopsis thalia
269	25	80.6	390	22	ABR71416	Drosophila melanog
270	25	80.6	391	23	ABR40035	Staphylococcus epi
271	25	80.6	400	22	AAW03593	Styphen tryptophan
272	25	80.6	402	22	ABR53504	Lactococcus lactis
273	25	80.6	410	22	AAW03592	Corn tryptophanyl-
274	25	80.6	418	22	ABR71839	Drosophila melanog
275	25	80.6	421	22	ABR67110	Drosophila melanog
276	25	80.6	432	22	AAW72396	Corn proline trans
277	25	80.6	436	21	AAW09653	Arabidopsis thalia
278	25	80.6	436	21	AAW36743	Arabidopsis thalia
279	25	80.6	436	21	AAW92008	Herbicidally activ
280	25	80.6	439	23	ABR92681	Herbicidally activ
281	25	80.6	442	21	AAW13571	Arabidopsis thalia
282	25	80.6	442	21	AAW17444	Arabidopsis thalia
283	25	80.6	442	21	AAW45471	Arabidopsis thalia
284	25	80.6	442	21	AAW45499	Arabidopsis thalia
285	25	80.6	442	23	ABR92053	Herbicidally activ
286	25	80.6	466	21	AAW30938	Arabidopsis thalia
287	25	80.6	467	21	AAW75499	Neisseria meningit
288	25	80.6	467	21	AAW27594	Neisseria meningit
289	25	80.6	472	21	AAW30937	Arabidopsis thalia
290	25	80.6	473	21	ABR82135	Rice proline trans
291	25	80.6	476	21	AAW35542	Human cap methyltr
292	25	80.6	476	21	AAW32245	Human cap methyltr
293	25	80.6	476	21	AAW13599	Human RNA (guanine
294	25	80.6	481	21	AAW30936	Arabidopsis thalia
295	25	80.6	481	23	AAW07510	Arabidopsis thalia
296	25	80.6	481	23	ABR93601	Herbicidally activ
297	25	80.6	487	19	AAW57484	Papaya ACC synthas
298	25	80.6	488	19	AAW98732	H. pylori GHPD 921
299	25	80.6	488	22	AAW35851	Helicobacter pylori
300	25	80.6	498	22	AAW082973	Human homologue of
301	25	80.6	505	22	ABR71008	Drosophila melanog
302	25	80.6	512	22	ABR61369	Drosophila melanog

Novel human diagno
Drosophila melanog
Drosophila melanog
Drosophila melanog
Human gene 38 enco
Human albumin Tusi
Gene 42 human secr
Staphylococcus epi
ORC3 subunit of Ye
S. cerevisiae orig
Human ORF2423
Human MAP/ERK kina
Drosophila melanog
Human ORF1887
Drosophila melanog
Novel human protei
Putative P. abyssi
A human dishevelle
Drosophila sp. PAR
Bacteriophage T7 R
Wild-type T7 RNA p
Wild-type T7 RNA p
Wild-type RNA poly
Mutant thermostabl
Modified bacterioph
A human dishevelle
Novel human diagno
P. falciparum 1-de
P. falciparum 1-de
Isoprenoid related
Ret brain serine e
Ret brain homologi
A. tetreus ORF1 es
HP1V-3 US isolate
HP1V-3 FRH1 cp45 v
HP1V-3 Vero cp45 v
Amino acid sequenc
Paraneoplastic cer
Rabbit albumin-bi
Peptide fragment c
Streptococcus pneu
Peptide #10662 enc
Human brain expres
Human bone marrow
Peptide #11029 enc
Human peptide enco
Cone snail alpha-c
Human immune/haema
Peptide #2 related
Amino acid sequenc
Propionibacterium
Human ORF1 protein
Drosophila melanog
Peptide #6805 enco
Protein #6111 enco
Human brain expres
Human bone marrow
Peptide #6060 enco
Peptide #6844 enco
Human peptide enco
Corynebacterium gl
Human ORF8 protei
Human brain expres
Human bone marrow
Human colon cancer
Human polypeptide
Galactol-1-phosp
Novel human diagno
Mouse germline hea
Human polypeptide
Human kinase-like
Staphylococcus epi
Bacteriophage lamb

376	24	77.4	115	21	AAV90814	33F8 hybridoma VH	449	24	77.4	233	22	AG91894	C glutamicum prote
377	24	77.4	116	17	AAU03642	Anti-adipocyte mon	450	24	77.4	233	23	AB949753	Listeria monocytog
378	24	77.4	116	22	AAU02515	Drosophila melanog	451	24	77.4	234	14	AA945009	Sequence encoded b
379	24	77.4	117	22	AB868972	Novel human diagno	452	24	77.4	234	23	AAU78086	Human CD30-11gang
380	24	77.4	118	22	ABG06010	Novel human diagno	453	24	77.4	236	14	AA944016	"Lys-63" E.coli he
381	24	77.4	119	18	AAW28320	Amino acid sequenc	454	24	77.4	236	14	AA944017	"Lys-97" E.coli he
382	24	77.4	119	22	ABG17655	Novel human diagno	455	24	77.4	236	14	AA944018	"Tyr-97" E.coli he
383	24	77.4	124	23	ABP03481	Human ORFX protein	456	24	77.4	236	14	AA944019	"Glu-107" E.coli h
384	24	77.4	125	16	AA880333	Protein polymERIC	457	24	77.4	236	14	AA944020	"Lys-104" E.coli h
385	24	77.4	128	22	AA887626	Bovine mammary tis	458	24	77.4	236	14	AA944021	"Asp-104" E.coli h
386	24	77.4	129	21	AA856935	Breast and ovarian	459	24	77.4	236	14	AA944022	"Ser-104" E.coli h
387	24	77.4	129	23	AB879472	Human TNFR/NGFR 14	460	24	77.4	236	14	AA944023	"Ser-106" E.coli h
388	24	77.4	130	21	AAAG01027	Human secreted pro	461	24	77.4	236	14	AA944024	"Glu-114" E.coli h
389	24	77.4	130	23	ABP06238	Human ORFX protein	462	24	77.4	236	14	AA944025	"Lys-114" E.coli h
390	24	77.4	132	21	AAAG01026	Human secreted pro	463	24	77.4	236	14	AA944026	E.coli heat labile
391	24	77.4	133	22	ABG03939	Novel human diagno	464	24	77.4	236	20	AA967772	Human protein sequ
392	24	77.4	133	22	ABG06007	Novel human diagno	465	24	77.4	237	20	AB807780	E.coli heat labil
393	24	77.4	138	22	AB856073	Skin cell protein,	466	24	77.4	237	20	AB807780	E.coli mutant heat
394	24	77.4	138	22	AB872273	Murine protein iso	467	24	77.4	238	20	AAU39950	Renilla mulleri gr
395	24	77.4	139	22	AA880720	Human haematologic	468	24	77.4	238	22	AAU08990	R. reniformis Gree
396	24	77.4	139	22	AA881194	Human haematologic	469	24	77.4	238	22	AA883915	Amino acid sequenc
397	24	77.4	139	22	AA881591	Human haematologic	470	24	77.4	238	22	ABP51820	Renilla reniformis
398	24	77.4	139	22	AA881830	Human haematologic	471	24	77.4	238	23	AAE13380	Renilla mulleri gr
399	24	77.4	141	21	AAV76072	Human skin cell pr	472	24	77.4	239	22	AAU08991	R. reniformis huma
400	24	77.4	141	22	AA856011	Skin cell protein,	473	24	77.4	239	22	ABP51821	Humanised Renilla
401	24	77.4	141	23	AB872211	Human protein isol	474	24	77.4	240	19	AAW65074	E.coli LT-A prote
402	24	77.4	142	23	AB853522	Lactococcus lactis	475	24	77.4	240	19	AAW65075	E.coli LT-A mutan
403	24	77.4	143	21	AA808276	Amino acid sequenc	476	24	77.4	240	23	AB807778	E.coli heat-labile
404	24	77.4	147	23	AB878638	Rat Ostr12 protein	477	24	77.4	242	18	AAW55535	H. pylori ORF 02ep
405	24	77.4	151	20	AAV34489	Porphorymonas ging	478	24	77.4	244	22	AAU40335	Propionibacterium
406	24	77.4	159	20	AAV34364	H. pylori CHPO 132	479	24	77.4	244	22	AAU50848	Propionibacterium
407	24	77.4	160	19	AAW68807	Staphylococcus epi	480	24	77.4	245	22	AB848557	Listeria monocytog
408	24	77.4	160	23	ABP39482	Novel human diagno	481	24	77.4	245	22	AB816388	Novel human diagno
409	24	77.4	161	23	ABG02526	Novel human diagno	482	24	77.4	245	22	AAW23392	Human EST encoded
410	24	77.4	168	22	ABG03881	Ubiquinone oxidore	483	24	77.4	245	22	AAW24000	Human EST encoded
411	24	77.4	172	21	AAV32425	Gene 2 human secre	484	24	77.4	246	22	ABG21287	Novel human diagno
412	24	77.4	176	21	AA845174	Human secreted pro	485	24	77.4	247	23	ABP45007	Human Blys binding
413	24	77.4	176	21	AA845175	Arabiidopsis thalia	486	24	77.4	247	23	ABP45195	Human Blys binding
414	24	77.4	177	17	AA898370	Mycobacterial AhpF	487	24	77.4	249	23	ABP44998	Human Blys binding
415	24	77.4	177	21	AA823822	Novel human diagno	488	24	77.4	249	23	ABP45008	Human Blys binding
416	24	77.4	184	22	ABG26613	Arabiidopsis thalia	489	24	77.4	249	23	ABP45013	Human Blys binding
417	24	77.4	185	23	ABP60704	Arabiidopsis thalia	490	24	77.4	249	23	ABP45112	Human Blys binding
418	24	77.4	186	21	AA853624	Human colon cancer	491	24	77.4	249	23	ABP45149	Human Blys binding
419	24	77.4	186	21	AA821658	Arabiidopsis thalia	492	24	77.4	249	23	ABP45191	Human Blys binding
420	24	77.4	187	22	AAE09736	Human kinase inter	493	24	77.4	249	23	ABP45366	Human Blys binding
421	24	77.4	190	21	AA809214	Arabiidopsis thalia	494	24	77.4	249	23	ABP45559	Human Blys binding
422	24	77.4	192	16	AA880345	Protein polymERIC	495	24	77.4	249	23	ABP45560	Human Blys binding
423	24	77.4	193	22	ABG26614	Novel human diagno	496	24	77.4	249	23	ABP45561	Human Blys binding
424	24	77.4	193	22	AA892454	Human protein sequ	497	24	77.4	249	23	ABP45562	Human Blys binding
425	24	77.4	194	21	AA809213	Herbicideally activ	498	24	77.4	250	23	ABP45003	Human Blys binding
426	24	77.4	194	23	AB892772	Protein polymERIC	499	24	77.4	250	23	ABP45148	Human Blys binding
427	24	77.4	198	16	AA880346	Protein polymERIC	500	24	77.4	250	23	ABP45157	Human Blys binding
428	24	77.4	198	19	AAW49722	Collagen-like poly	501	24	77.4	250	23	ABP45201	Human Blys binding
429	24	77.4	198	19	AAW57672	Streptococcus poly	502	24	77.4	250	23	ABP45524	Human Blys binding
430	24	77.4	200	23	ABP26704	Arabiidopsis thalia	503	24	77.4	250	23	ABP45529	Human Blys binding
431	24	77.4	201	21	AA816880	Staphylococcus epi	504	24	77.4	250	23	ABP45552	Human Blys binding
432	24	77.4	203	23	ABP38851	Arabiidopsis thalia	505	24	77.4	250	23	ABP45553	Human Blys binding
433	24	77.4	204	21	AA823821	Human cancer assoc	506	24	77.4	250	23	ABP45554	Human Blys binding
434	24	77.4	209	21	AA843593	Human cancer assoc	507	24	77.4	250	23	ABP45555	Human Blys binding
435	24	77.4	209	21	AA899984	Chick limb deforma	508	24	77.4	250	23	ABP45563	Human Blys binding
436	24	77.4	213	21	AA826194	Human mitotin/kine	509	24	77.4	250	23	ABP45566	Human Blys binding
437	24	77.4	215	14	AA845007	Sequence encoded b	510	24	77.4	251	19	AAW49735	Protein polymer ad
438	24	77.4	216	18	AAW28267	Staphylococcus aur	511	24	77.4	251	23	ABP44879	Human Blys binding
439	24	77.4	218	23	AAE23568	Chick heat shock c	512	24	77.4	251	23	ABP44880	Human Blys binding
440	24	77.4	218	23	AAE23593	Drosophila pseudoo	513	24	77.4	251	23	ABP44883	Human Blys binding
441	24	77.4	218	23	AAE23594	Cricetulus griseus	514	24	77.4	251	23	ABP44892	Human Blys binding
442	24	77.4	218	23	AAE23596	Drosophila melanog	515	24	77.4	251	23	ABP44899	Human Blys binding
443	24	77.4	218	23	AAE23597	Human heat shock p	516	24	77.4	251	23	ABP44900	Human Blys binding
444	24	77.4	218	23	AAE23598	Chick heat shock p	517	24	77.4	251	23	ABP45000	Human Blys binding
445	24	77.4	218	23	AAE23612	Rattus species hea	518	24	77.4	251	23	ABP45002	Human Blys binding
446	24	77.4	225	21	AA845121	Human secreted pro	519	24	77.4	251	23	ABP45009	Human Blys binding
447	24	77.4	230	22	AA876724	Corynebacterium q1	520	24	77.4	251	23	ABP45110	Human Blys binding
448	24	77.4	231	22	AA868745	Human immune/haema	521	24	77.4	251	23	ABP45111	Human Blys binding

522	24	77.4	251	23	ABP45115	Human	Blvs binding	595	24	77.4	259	21	AAV96645	Plant-optimized E.
523	24	77.4	251	23	ABP45116	Human	Blvs binding	596	24	77.4	259	21	AAV96647	Synthetic E. coli
524	24	77.4	251	23	ABP45119	Human	Blvs binding	597	24	77.4	259	21	AAV96648	Plant-optimized E.
525	24	77.4	251	23	ABP45129	Human	Blvs binding	598	24	77.4	259	21	AAV96649	Plant-optimized E.
526	24	77.4	251	23	ABP45130	Human	Blvs binding	599	24	77.4	259	21	AAV96650	Plant-optimized E.
527	24	77.4	251	23	ABP45137	Human	Blvs binding	600	24	77.4	259	21	AAV96651	Plant-optimized E.
528	24	77.4	251	23	ABP45137	Human	Blvs binding	601	24	77.4	260	21	AAV96652	Arabidopsis thalia
529	24	77.4	251	23	ABP45137	Human	Blvs binding	602	24	77.4	261	18	AAW25125	Single chain pro-r
530	24	77.4	251	23	ABP45140	Human	Blvs binding	603	24	77.4	261	18	AAW21712	RIP fusion protein
531	24	77.4	251	23	ABP45141	Human	Blvs binding	604	24	77.4	262	19	AAW49734	Protein polymer ad
532	24	77.4	251	23	ABP45144	Human	Blvs binding	605	24	77.4	265	18	AAW25131	Pro-ribosome inact
533	24	77.4	251	23	ABP45146	Human	Blvs binding	606	24	77.4	265	18	AAW21718	RIP fusion protein
534	24	77.4	251	23	ABP45151	Human	Blvs binding	607	24	77.4	269	20	AAW87568	Plasmid pUK21-A Al
535	24	77.4	251	23	ABP45155	Human	Blvs binding	608	24	77.4	271	13	AAW20993	Protein encoded by
536	24	77.4	251	23	ABP45200	Human	Blvs binding	609	24	77.4	271	14	AAW34542	kan gene product f
537	24	77.4	251	23	ABP45205	Human	Blvs binding	610	24	77.4	271	14	AAW14836	Protein encoded by
538	24	77.4	251	23	ABP45208	Human	Blvs binding	611	24	77.4	271	19	AAW56597	kanamycin resist
539	24	77.4	251	23	ABP45209	Human	Blvs binding	612	24	77.4	271	20	AAV42547	Plasmid pRZT1, ka
540	24	77.4	251	23	ABP45213	Human	Blvs binding	613	24	77.4	271	20	AAV15385	Plasmid pRZT1, ka
541	24	77.4	251	23	ABP45221	Human	Blvs binding	614	24	77.4	271	20	AAV07249	Bacterial kanamyci
542	24	77.4	251	23	ABP45362	Human	Blvs binding	615	24	77.4	271	21	AAV70006	Protein product of
543	24	77.4	251	23	ABP45490	Human	Blvs binding	616	24	77.4	271	22	AAW85258	Protein product of
544	24	77.4	251	23	ABP45545	Human	Blvs binding	617	24	77.4	271	22	AAW85258	Protein product of
545	24	77.4	251	23	ABP45580	Human	Blvs binding	618	24	77.4	271	22	AAW85258	Protein product of
546	24	77.4	251	23	ABP45701	Human	Blvs binding	619	24	77.4	271	22	AAW85258	Protein product of
547	24	77.4	251	23	ABP45821	Human	Blvs binding	620	24	77.4	271	22	AAW85258	Protein product of
548	24	77.4	251	23	ABP45826	Human	Blvs binding	621	24	77.4	271	22	AAW85258	Protein product of
549	24	77.4	251	23	ABP45829	Human	Blvs binding	622	24	77.4	271	22	AAW85258	Protein product of
550	24	77.4	251	23	ABP45886	Human	Blvs binding	623	24	77.4	271	22	AAW85258	Protein product of
551	24	77.4	251	23	ABP45905	Human	Blvs binding	624	24	77.4	271	22	AAW85258	Protein product of
552	24	77.4	251	23	ABP45910	Human	Blvs binding	625	24	77.4	271	22	AAW85258	Protein product of
553	24	77.4	251	23	ABP45937	Human	Blvs binding	626	24	77.4	271	22	AAW85258	Protein product of
554	24	77.4	251	23	ABP45944	Human	Blvs binding	627	24	77.4	271	22	AAW85258	Protein product of
555	24	77.4	251	23	ABP45944	Human	Blvs binding	628	24	77.4	271	22	AAW85258	Protein product of
556	24	77.4	251	23	ABP45944	Human	Blvs binding	629	24	77.4	271	22	AAW85258	Protein product of
557	24	77.4	251	23	ABP45944	Human	Blvs binding	630	24	77.4	271	22	AAW85258	Protein product of
558	24	77.4	251	23	ABP45944	Human	Blvs binding	631	24	77.4	271	22	AAW85258	Protein product of
559	24	77.4	251	23	ABP45944	Human	Blvs binding	632	24	77.4	271	22	AAW85258	Protein product of
560	24	77.4	251	23	ABP45944	Human	Blvs binding	633	24	77.4	271	22	AAW85258	Protein product of
561	24	77.4	251	23	ABP45944	Human	Blvs binding	634	24	77.4	271	22	AAW85258	Protein product of
562	24	77.4	251	23	ABP45944	Human	Blvs binding	635	24	77.4	271	22	AAW85258	Protein product of
563	24	77.4	251	23	ABP45944	Human	Blvs binding	636	24	77.4	271	22	AAW85258	Protein product of
564	24	77.4	251	23	ABP45944	Human	Blvs binding	637	24	77.4	271	22	AAW85258	Protein product of
565	24	77.4	251	23	ABP45944	Human	Blvs binding	638	24	77.4	271	22	AAW85258	Protein product of
566	24	77.4	251	23	ABP45944	Human	Blvs binding	639	24	77.4	271	22	AAW85258	Protein product of
567	24	77.4	251	23	ABP45944	Human	Blvs binding	640	24	77.4	271	22	AAW85258	Protein product of
568	24	77.4	251	23	ABP45944	Human	Blvs binding	641	24	77.4	271	22	AAW85258	Protein product of
569	24	77.4	251	23	ABP45944	Human	Blvs binding	642	24	77.4	271	22	AAW85258	Protein product of
570	24	77.4	251	23	ABP45944	Human	Blvs binding	643	24	77.4	271	22	AAW85258	Protein product of
571	24	77.4	251	23	ABP45944	Human	Blvs binding	644	24	77.4	271	22	AAW85258	Protein product of
572	24	77.4	251	23	ABP45944	Human	Blvs binding	645	24	77.4	271	22	AAW85258	Protein product of
573	24	77.4	251	23	ABP45944	Human	Blvs binding	646	24	77.4	271	22	AAW85258	Protein product of
574	24	77.4	251	23	ABP45944	Human	Blvs binding	647	24	77.4	271	22	AAW85258	Protein product of
575	24	77.4	251	23	ABP45944	Human	Blvs binding	648	24	77.4	271	22	AAW85258	Protein product of
576	24	77.4	251	23	ABP45944	Human	Blvs binding	649	24	77.4	271	22	AAW85258	Protein product of
577	24	77.4	251	23	ABP45944	Human	Blvs binding	650	24	77.4	271	22	AAW85258	Protein product of
578	24	77.4	251	23	ABP45944	Human	Blvs binding	651	24	77.4	271	22	AAW85258	Protein product of
579	24	77.4	251	23	ABP45944	Human	Blvs binding	652	24	77.4	271	22	AAW85258	Protein product of
580	24	77.4	251	23	ABP45944	Human	Blvs binding	653	24	77.4	271	22	AAW85258	Protein product of
581	24	77.4	251	23	ABP45944	Human	Blvs binding	654	24	77.4	271	22	AAW85258	Protein product of
582	24	77.4	251	23	ABP45944	Human	Blvs binding	655	24	77.4	271	22	AAW85258	Protein product of
583	24	77.4	251	23	ABP45944	Human	Blvs binding	656	24	77.4	271	22	AAW85258	Protein product of
584	24	77.4	251	23	ABP45944	Human	Blvs binding	657	24	77.4	271	22	AAW85258	Protein product of
585	24	77.4	251	23	ABP45944	Human	Blvs binding	658	24	77.4	271	22	AAW85258	Protein product of
586	24	77.4	251	23	ABP45944	Human	Blvs binding	659	24	77.4	271	22	AAW85258	Protein product of
587	24	77.4	251	23	ABP45944	Human	Blvs binding	660	24	77.4	271	22	AAW85258	Protein product of
588	24	77.4	251	23	ABP45944	Human	Blvs binding	661	24	77.4	271	22	AAW85258	Protein product of
589	24	77.4	251	23	ABP45944	Human	Blvs binding	662	24	77.4	271	22	AAW85258	Protein product of
590	24	77.4	251	23	ABP45944	Human	Blvs binding	663	24	77.4	271	22	AAW85258	Protein product of
591	24	77.4	251	23	ABP45944	Human	Blvs binding	664	24	77.4	271	22	AAW85258	Protein product of
592	24	77.4	251	23	ABP45944	Human	Blvs binding	665	24	77.4	271	22	AAW85258	Protein product of
593	24	77.4	251	23	ABP45944	Human	Blvs binding	666	24	77.4	271	22	AAW85258	Protein product of
594	24	77.4	251	23	ABP45944	Human	Blvs binding	667	24	77.4	271	22	AAW85258	Protein product of

668	24	77.4	315	21	AA68460	Sphingomonas sp. C	741	24	77.4	398	23	AA021885	Isoprenoid related
669	24	77.4	316	22	AA67156	M. crenulata K1H2	742	24	77.4	402	13	AA620084	Maize RIP deriv. R
670	24	77.4	317	21	AA610986	M. crenulata Hemoc	743	24	77.4	402	18	AA625129	Pro-rhosome inact
671	24	77.4	317	22	AA671098	M. crenulata K1H2	744	24	77.4	402	18	AA621716	RIP fusion protein
672	24	77.4	318	22	AA640912	Human polypeptide	745	24	77.4	402	20	AA625164	Maize K1P protein
673	24	77.4	320	20	AA639288	Human phosphodiester	746	24	77.4	404	23	AA628391	Streptococcus poly
674	24	77.4	320	21	AA645743	Arabidopsis thalia	747	24	77.4	405	22	AA632963	C glutamicum prote
675	24	77.4	323	21	AA620858	Arabidopsis thalia	748	24	77.4	412	21	AA649547	Arabidopsis thalia
676	24	77.4	323	23	AA661055	Lactobacillus tham	749	24	77.4	412	21	AA695049	Arabidopsis thalia
677	24	77.4	325	20	AA697705	Staphylococcus aur	750	24	77.4	412	22	AA655334	Human protein sequ
678	24	77.4	327	21	AA616879	Arabidopsis thalia	751	24	77.4	415	21	AA651603	Arabidopsis thalia
679	24	77.4	327	22	AA605604	Novel human diago	752	24	77.4	419	22	AA656269	putative P. abysci
680	24	77.4	331	17	AA681570	Transglutaminase.	753	24	77.4	419	21	AA651602	Arabidopsis thalia
681	24	77.4	331	21	AA621657	Arabidopsis thalia	754	24	77.4	420	21	AA634980	Arabidopsis thalia
682	24	77.4	331	23	AA606745	Streptomyces lydic	755	24	77.4	420	22	AA679310	Arabidopsis thalia
683	24	77.4	336	21	AA645742	Arabidopsis thalia	756	24	77.4	420	23	AA680836	Corynebacterium q1
684	24	77.4	336	22	AA602575	Enterococcus typt	757	24	77.4	422	22	AA607536	Corynebacterium q1
685	24	77.4	337	21	AA616878	Arabidopsis thalia	758	24	77.4	422	22	AA607536	Novel human diago
686	24	77.4	337	21	AA645741	Arabidopsis thalia	759	24	77.4	423	21	AA653325	Novel human diago
687	24	77.4	338	16	AA680350	Protein polymeric	760	24	77.4	424	21	AA630363	Arabidopsis thalia
688	24	77.4	339	21	AA620857	Arabidopsis thalia	761	24	77.4	425	22	AA634976	Maize K1N17 orthol
689	24	77.4	340	21	AA620856	Arabidopsis thalia	762	24	77.4	426	22	AA634976	Enterococcus faeca
690	24	77.4	341	21	AA621656	Arabidopsis thalia	763	24	77.4	430	13	AA626173	Part of Major yo p
691	24	77.4	346	22	AA606006	Novel human diago	764	24	77.4	430	20	AA639286	Phosphodiesterase
692	24	77.4	347	22	AA636873	Staphylococcus aur	765	24	77.4	430	22	AA665234	Human leukocyte al
693	24	77.4	348	22	AA607777	Novel human diago	766	24	77.4	431	23	AA697375	Novel human protei
694	24	77.4	348	22	AA608483	Novel human diago	767	24	77.4	432	22	AA660787	Drosophila melanog
695	24	77.4	350	21	AA617409	Arabidopsis thalia	768	24	77.4	433	21	AA633095	Arabidopsis thalia
696	24	77.4	350	23	AA691589	Herbicidially activ	769	24	77.4	433	23	AA677386	Myxococcus acyl Co
697	24	77.4	351	23	AA690981	Herbicidially activ	770	24	77.4	435	21	AA640644	Zea mays protein f
698	24	77.4	351	23	AA693636	Herbicidially activ	771	24	77.4	435	21	AA640937	Zea mays protein f
699	24	77.4	352	22	AA693427	Human protein sequ	772	24	77.4	436	21	AA640643	Zea mays protein f
700	24	77.4	352	15	AA615384	Commamonas testoste	773	24	77.4	436	22	AA64829	Human protein sequ
701	24	77.4	354	20	AA697714	Staphylococcus aur	774	24	77.4	436	22	AA64829	Herbicidially activ
702	24	77.4	356	23	AA675492	S. aureus antigen	775	24	77.4	438	23	AA692558	Herbicidially activ
703	24	77.4	358	22	AA681462	S. epidermidis ope	776	24	77.4	442	22	AA603316	Novel human diago
704	24	77.4	359	19	AA698324	H. pylori GHP0 141	777	24	77.4	443	21	AA649136	Arabidopsis thalia
705	24	77.4	362	13	AA698324	Staphylococcus epi	778	24	77.4	449	23	AA692561	Herbicidially activ
706	24	77.4	363	18	AA698789	Staphylococcus aur	779	24	77.4	449	23	AA693864	Herbicidially activ
707	24	77.4	367	17	AA694501	Polyhydroxyacid sy	780	24	77.4	450	22	AA648720	Listeria monocytog
708	24	77.4	368	12	AA694501	Human BPPI/Kuntiz	781	24	77.4	452	23	AA656342	Propionibacterium
709	24	77.4	370	23	AA607785	E coli mutant heat	782	24	77.4	452	23	AA692562	Herbicidially activ
710	24	77.4	373	23	AA607784	E coli heat-labile	783	24	77.4	453	23	AA693638	Herbicidially activ
711	24	77.4	375	23	AA697201	Novel human protei	784	24	77.4	456	21	AA640642	Zea mays protein f
712	24	77.4	378	17	AA609219	SELPK polymer. S	785	24	77.4	458	22	AA623932	Human carboxypepti
713	24	77.4	378	19	AA653545	Amino acid sequenc	786	24	77.4	458	22	AA624116	Human EST encoded
714	24	77.4	378	21	AA651889	Plasmid pPT0375 pr	787	24	77.4	458	23	AA693883	Herbicidially activ
715	24	77.4	380	22	AA600507	E. coli heat-labil	788	24	77.4	460	22	AA630856	Amino acid sequenc
716	24	77.4	382	22	AA600506	E. coli heat-labil	789	24	77.4	460	22	AA630856	Zea mays protein f
717	24	77.4	382	22	AA693621	Human protein sequ	790	24	77.4	463	21	AA640935	Human wild type g1
718	24	77.4	385	21	AA625461	Arabidopsis thalia	791	24	77.4	464	14	AA650197	Human liver glucoc
719	24	77.4	385	21	AA634981	Arabidopsis thalia	792	24	77.4	464	23	AA622738	Human wild type g1
720	24	77.4	385	23	AA692563	Herbicidially activ	793	24	77.4	465	14	AA685019	Human novel cytoxi
721	24	77.4	386	22	AA628045	Novel human secret	794	24	77.4	465	22	AA692790	Human protein sequ
722	24	77.4	387	20	AA625163	Maize RIP-5 protei	795	24	77.4	466	14	AA650196	Human wild type g1
723	24	77.4	388	21	AA651604	Arabidopsis thalia.	796	24	77.4	466	20	AA639283	Phosphodiesterase
724	24	77.4	390	23	AA635303	Human familial com	797	24	77.4	469	21	AA649135	Arabidopsis thalia
725	24	77.4	391	20	AA625619	Murine kin17 prote	798	24	77.4	472	12	AA614196	Human cannabinoi
726	24	77.4	391	20	AA660850	Mevlonate pathway	799	24	77.4	472	12	AA656338	Non-endorogenous hum
727	24	77.4	391	23	AA678132	Human familial com	800	24	77.4	473	12	AA614195	Rat cannabinoi
728	24	77.4	391	23	AA678132	Human VDP-1 prote	801	24	77.4	473	23	AA657048	Mouse ischaemic co
729	24	77.4	393	20	AA678233	A human kin17 prot	802	24	77.4	474	22	AA692673	Human protein sequ
730	24	77.4	395	23	AA635302	Mouse HYP1P1. Ma	803	24	77.4	474	22	AA692673	Drosophila melanog
731	24	77.4	395	23	AA678131	Mouse HYP1P1. Ma	804	24	77.4	475	22	AA663526	Novel human secret
732	24	77.4	395	23	AA648926	Regulatory protein	805	24	77.4	478	22	AA630550	Protein polymer ad
733	24	77.4	396	15	AA660777	Mythamycin IV hyd	806	24	77.4	479	19	AA649739	Lactococcus lactis
734	24	77.4	397	21	AA653326	Arabidopsis thalia	807	24	77.4	482	22	AA686967	D. melanogaster pe
735	24	77.4	398	20	AA652838	Escherichia coli p	808	24	77.4	482	21	AA610901	S. xylois D1A pr
736	24	77.4	398	21	AA658134	Lung cancer associ	809	24	77.4	487	21	AA633456	Human PRO1917 prot
737	24	77.4	398	21	AA623096	Arabidopsis thalia	810	24	77.4	487	21	AA656734	PRO1917, an inosit
738	24	77.4	398	22	AA637784	Streptococcus pneu	811	24	77.4	487	22	AA612435	Human PRO1917 poly
739	24	77.4	398	22	AA601066	CPE 69 protein seq	812	24	77.4	487	22	AA650929	Human PRO1917 prot
740	24	77.4	398	23	AA621878	Isoprenoid related	813	24	77.4	487	22	AA643383	Amino acid sequenc

814	24	77.4	487	23	ABB95547	Human angiogenesis
815	24	77.4	487	23	ABB84841	Human PRO1917 prot
816	24	77.4	487	23	AAU083693	Human PRO protein,
817	24	77.4	489	13	AAAR20085	Matze RIP deriv. R
818	24	77.4	489	18	AAAW25130	Pro-ribosome inact
819	24	77.4	489	18	AAAW21717	RIP fusion protein
820	24	77.4	490	20	AAAY13366	Amino acid sequenc
821	24	77.4	490	22	AAAU12326	Human PRO222 polyp
822	24	77.4	490	22	AAAB80234	Human PRO222 prote
823	24	77.4	492	23	AAAB65035	Human albumin fusi
824	24	77.4	492	23	AAAU91358	Human secreted pro
825	24	77.4	492	23	AAU074621	Oestrogen-regulate
826	24	77.4	493	18	AAAW25132	Pro-ribosome inact
827	24	77.4	493	18	AAAW21719	RIP fusion protein
828	24	77.4	494	22	AAAB62540	Drosophila melanog
829	24	77.4	501	22	AAAB58136	Drosophila melanog
830	24	77.4	502	18	AAAW25133	Pro-ribosome inact
831	24	77.4	502	18	AAAW21720	RIP fusion protein
832	24	77.4	508	22	AAAB59820	Drosophila melanog
833	24	77.4	509	13	AAAR30429	Major yo paraneopl
834	24	77.4	514	20	AAAY33675	Human DKL1 protein
835	24	77.4	514	22	AAAB90586	Human secreted pro
836	24	77.4	517	22	AAAB65472	Human albumin fusi
837	24	77.4	517	22	AAAE05394	Staphylococcus aur
838	24	77.4	520	22	AAU034239	Staphylococcus aur
839	24	77.4	520	22	AAAG90708	C glutamicum prote
840	24	77.4	520	22	AAAB80105	Cornebacterium gl
841	24	77.4	521	22	AAAP05392	Staphylococcus aur
842	24	77.4	521	23	AAAS50327	Human nucleotide b
843	24	77.4	531	20	AAAY05376	Human HCW Inducib
844	24	77.4	532	17	AAAR91035	Recombinant V8 pro
845	24	77.4	532	18	AAAW22219	Protein encoded by
846	24	77.4	533	20	AAAY39284	Phosphodiesterase
847	24	77.4	533	22	AAAB92829	Human protein sequ
848	24	77.4	536	21	AAAG49134	Arbidiopsis thalia
849	24	77.4	537	18	AAAW22220	Protein encoded by
850	24	77.4	537	22	AAAB60496	Human cell cycle a
851	24	77.4	537	23	AAU074622	Oestrogen-regulate
852	24	77.4	540	22	AAU030977	Novel human secret
853	24	77.4	544	20	AAAY14049	G. oxydans D-sorbi
854	24	77.4	544	20	AAAY14049	G. oxydans D-sorbi
855	24	77.4	555	16	AAAR77864	S. clavuligerus OR
856	24	77.4	556	22	AAAE07912	S. clavuligerus cl
857	24	77.4	559	22	AAAB26803	Novel human diagno
858	24	77.4	561	22	AAAB58264	Drosophila melanog
859	24	77.4	562	22	AAU038523	Drosophila G-prote
860	24	77.4	562	23	AAU033507	Enterococcus faeca
861	24	77.4	569	23	AAAB50001	Listeria monocytog
862	24	77.4	572	22	AAAU5140	Enterococcus faeca
863	24	77.4	580	17	AAAM02164	Lactococcus lactis
864	24	77.4	580	20	AAAY39285	Phosphodiesterase
865	24	77.4	587	21	AAAG49546	Arbidiopsis thalia
866	24	77.4	593	20	AAAY28561	Cyclic-GMP specifi
867	24	77.4	593	23	AAAB61846	Prostate cancer-as
868	24	77.4	595	22	AAU049951	Propionibacterium
869	24	77.4	600	21	AAAG49545	Arbidiopsis thalia
870	24	77.4	600	22	AAAU64743	Propionibacterium
871	24	77.4	603	22	AAAG82295	S. epidermidis ope
872	24	77.4	603	23	AAAB40381	Staphylococcus epi
873	24	77.4	606	23	AAU017858	Pyrim domain conta
874	24	77.4	607	21	AAAG53324	Arbidiopsis thalia
875	24	77.4	608	22	AAAB99139	Pyrococcus abyssi
876	24	77.4	608	22	AAAU16168	Human novel secret
877	24	77.4	611	22	AAAU12809	Soybean arginyl-tr
878	24	77.4	613	18	AAAW22051	DNA polymerase 220
879	24	77.4	613	21	AAAU97097	Pfu DNA polymerase
880	24	77.4	613	21	AAAY52020	P. fuitosus PFUORF
881	24	77.4	615	17	AAAY51649	Pyrococcus furiosu
882	24	77.4	615	21	AAAM00456	1-Fructan
883	24	77.4	616	21	AAAG40603	Arbidiopsis thalia
884	24	77.4	617	21	AAAR24788	Plant SDF encoded
885	24	77.4	618	21	AAAG40602	Arbidiopsis thalia
886	24	77.4	629	22	AAAB73229	Human phosphatase
	24	77.4	631	22	AAAB58778	Drosophila melanog
887	24	77.4	887	21	AAAB24787	Plant SDF encoded
888	24	77.4	888	21	AAAG40601	Arbidiopsis thalia
889	24	77.4	889	23	AAAB51861	Lactococcus lactis
890	24	77.4	890	23	AAAB47300	Listeria monocytog
891	24	77.4	891	24	AAAB24786	Plant SDF encoded
892	24	77.4	892	16	AAAR83249	Chloroperoxidase.
893	24	77.4	893	24	AAAG00626	Novel human diagno
894	24	77.4	894	24	AAAB51288	Drosophila melanog
895	24	77.4	895	22	AAAG43584	Arbidiopsis thalia
896	24	77.4	896	21	AAAB42482	Human ORFX ORF2246
897	24	77.4	897	24	AAAG42583	Arbidiopsis thalia
898	24	77.4	898	21	AAAB03758	Human H37 amino ac
899	24	77.4	899	24	AAAG10004	Novel human diagno
900	24	77.4	900	23	AAAG6069	N. meningitidis re
901	24	77.4	901	24	AAAG31900	Arbidiopsis thalia
902	24	77.4	902	16	AAAR80330	Protein polymeric
903	24	77.4	903	19	AAAW45717	Collagen-Like poly
904	24	77.4	904	24	AAAW57670	Plant SDF encoded
905	24	77.4	905	21	AAAB24850	Arbidiopsis thalia
906	24	77.4	906	24	AAAG31899	Human secreted pro
907	24	77.4	907	24	AAAB11820	Drosophila melanog
908	24	77.4	908	22	AAAB58231	Drosophila melanog
909	24	77.4	909	22	AAAB67404	Drosophila melanog
910	24	77.4	910	22	AAAB93318	Herbicideally activ
911	24	77.4	911	22	AAAB59166	Drosophila melanog
912	24	77.4	912	22	AAAB66052	Drosophila melanog
913	24	77.4	913	22	AAAG98345	Escherichia coli p
914	24	77.4	914	22	AAAB24849	Plant SDF encoded
915	24	77.4	915	21	AAAB24848	Plant SDF encoded
916	24	77.4	916	22	AAAB61081	Drosophila melanog
917	24	77.4	917	22	AAAB57943	Drosophila melanog
918	24	77.4	918	22	AAAB36505	Drosophila Hsp83 p
919	24	77.4	919	22	AAAE12989	Human Hsp90 family
920	24	77.4	920	22	AAAB82537	Human heat shock p
921	24	77.4	921	22	AAAB36507	Human Hsp90 beta p
922	24	77.4	922	23	AAAB06994	Human Hsp90 beta p
923	24	77.4	923	12	AAAM05381	Chicken heat shock
924	24	77.4	924	17	AAAR10426	Human testicular a
925	24	77.4	925	21	AAAG42582	Arbidiopsis thalia
926	24	77.4	926	22	AAAE12988	Human Hsp90 family
927	24	77.4	927	22	AAAB82536	Human heat shock p
928	24	77.4	928	23	AAU071814	Human antigen Huhs
929	24	77.4	929	22	AAAM39446	Human polypeptide
930	24	77.4	930	22	AAAG21289	Novel human diagno
931	24	77.4	931	22	AAAM39447	Human polypeptide
932	24	77.4	932	22	AAAB84252	Human macro protei
933	24	77.4	933	16	AAAR80327	Protein polymeric
934	24	77.4	934	16	AAAR80324	Protein polymeric
935	24	77.4	935	19	AAAW49713	Protein polymer ad
936	24	77.4	936	19	AAAW49715	Protein polymer ad
937	24	77.4	937	19	AAAW57666	Collagen-Like poly
938	24	77.4	938	19	AAAW57668	Collagen-Like poly
939	24	77.4	939	22	AAAG53535	Amino acid sequenc
940	24	77.4	940	19	AAAG05609	Novel human diagno
941	24	77.4	941	13	AAAB97543	Novel human protei
942	24	77.4	942	12	AAAR15143	Hemochus contort
943	24	77.4	943	18	AAAW36508	Human RENT1 protei
944	24	77.4	944	24	AAAM41233	Human polypeptide
945	24	77.4	945	22	AAAM41234	Human polypeptide
946	24	77.4	946	22	AAU09864	Novel human secret
947	24	77.4	947	21	AAAG09864	Arbidiopsis thalia
948	24	77.4	948	21	AAAG31898	Drosophila melanog
949	24	77.4	949	22	AAAB65317	Drosophila melanog
950	24	77.4	950	22	AAAB65774	Streptococcus pyog
951	24	77.4	951	22	AAU03139	Streptococcus poly
952	24	77.4	952	16	AAAB26368	Streptococcus poly
953	24	77.4	953	16	AAAB80341	Streptococcus poly
954	24	77.4	954	19	AAAM09213	SEIPEK polymer. S
955	24	77.4	955	17	AAAM49728	SEIPEK polymer. S
956	24	77.4	956	24	AAAW53541	Expected amino aci
957	24	77.4	957	21	AAAY51882	Plasmid pPR0345 pr
958	24	77.4	958	22	AAU61465	Propionibacterium
959	24	77.4	959	22	AAU033722	Pseudomonas aerugi
	24	77.4	923	22	AAAG13717	Novel human diagno

960	24	77.4	936	17	AAW09221	SELPK-CS1 polymer
961	24	77.4	936	21	AAW51891	Plasmid pPT0370 pr
962	24	77.4	937	19	AAW53547	Amino acid sequenc
963	24	77.4	946	21	AAW29100	Human Inter-alpha-
964	24	77.4	966	17	AAW09223	SELPK-CS2 polymer
965	24	77.4	966	21	AAW51893	Plasmid pPT0373 cr
966	24	77.4	982	20	AAW55955	Nematode STE20-rel
967	24	77.4	1002	17	AAW09218	SELPK polymer. S
968	24	77.4	1002	19	AAW53544	Amino acid sequenc
969	24	77.4	1002	21	AAW51888	Plasmid pPT0364 pr
970	24	77.4	1007	22	AAW62079	Drosophila melanog
971	24	77.4	1032	23	AAW021588	Murine Toll-like r
972	24	77.4	1043	18	AAW36509	Murine RENT1 prote
973	24	77.4	1113	22	ABG10005	Novel human diagno
974	24	77.4	1115	22	AAW96812	putative P. abysal
975	24	77.4	1140	19	AAW21377	human HRP-1 wild
976	24	77.4	1145	22	ABW62531	Drosophila melanog
977	24	77.4	1157	22	AAW99502	CLASP-3 protein fr
978	24	77.4	1176	17	AAW88116	CryE(c) protein.
979	24	77.4	1177	22	AAW88117	CryE(d) protein.
980	24	77.4	1177	22	AAW64406	Amino acid sequenc
981	24	77.4	1184	23	AAW91131	Herbicideally activ
982	24	77.4	1194	21	AAW36529	Preliminary CLASP-
983	24	77.4	1194	21	AAW36530	Preliminary CLASP-
984	24	77.4	1194	21	AAW36531	Preliminary CLASP-
985	24	77.4	1194	21	AAW36532	Preliminary CLASP-
986	24	77.4	1194	21	AAW36533	Preliminary CLASP-
987	24	77.4	1194	21	AAW36534	Preliminary CLASP-
988	24	77.4	1194	21	AAW36535	Preliminary CLASP-
989	24	77.4	1194	21	AAW36536	Preliminary CLASP-
990	24	77.4	1194	22	AAW99503	CLASP-3 protein fr
991	24	77.4	1194	23	ABG61694	Cadherin-like asym
992	24	77.4	1194	23	ABG61695	Cadherin-like asym
993	24	77.4	1194	23	ABG61696	Cadherin-like asym
994	24	77.4	1194	23	ABG61697	Cadherin-like asym
995	24	77.4	1194	23	ABG61698	Cadherin-like asym
996	24	77.4	1194	23	ABG61699	Cadherin-like asym
997	24	77.4	1194	23	ABG61700	Cadherin-like asym
998	24	77.4	1194	23	ABG61701	Cadherin-like asym
999	24	77.4	1207	22	ABG06651	Novel human diagno
1000	24	77.4	1207	22	ABG07248	Novel human diagno

ALIGNMENTS

RESULT 1
ID ABB81561 standard; Protein: 169 AA.
XX
AC ABB81561;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human N-acetylglucosamine-6-sulfotransferase portion SEQ ID NO:9.
XX
KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme: GLCNAC6ST;
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological.
XX
OS Homo sapiens.
XX
PN US2002061562-A1.
XX
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU/) FUKUDA M N.
PA (AKAM/) AKAMA T O.
XX

FI Fukuda MN, Akama TO;
XX
DR WPI; 2002-507643/54.
XX
PT New nucleic acid encoding corneal
PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
PT monitoring and diagnosis of macular corneal dystrophy
XX
XX Example 2; Fig 3C; 69pp; English.
XX
XX The present invention describes human corneal
XX N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
XX sulfation of keratan sulfate (KS). Also described is a method for
XX monitoring the effect of treatments for macular corneal dystrophy (MCD),
XX and detecting susceptibility to MCD. (I) is located to chromosome 10q22,
XX and has ophthalmological activity. (I) can be used to treat or prevent
XX macular corneal dystrophy types I or II. (I) makes possible treatment
XX of MCD without requiring keratoplasty or keratectomy. The present
XX sequence represents a portion of a human N-acetylglucosamine-6-
XX sulfotransferase (GLCNAC6ST), which is given in the exemplification
XX of the present invention.
SQ Sequence 169 AA;

Query Match 100.0%; Score 31; DB 23; Length 169;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 127 VRYEDL 132

RESULT 2
ID ABB81560
XX ABB81560 standard; Protein: 171 AA.
XX
AC ABB81560;
XX
DT 05-SEP-2002 (first entry)
XX
DE Human high endothelial cell GLCNAC6ST portion SEQ ID NO:8.
XX
XX
KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme: GLCNAC6ST;
KW corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
KW ophthalmological.
XX
OS Homo sapiens.
XX
PN US2002061562-A1.
XX
PD 23-MAY-2002.
XX
PF 09-AUG-2001; 2001US-0927602.
XX
PR 11-AUG-2000; 2000US-325773P.
XX
PA (FUKU/) FUKUDA M N.
PA (AKAM/) AKAMA T O.
XX
XX Fukuda MN, Akama TO;
XX
XX WPI; 2002-507643/54.
XX
XX
XX New nucleic acid encoding corneal
XX N-acetylglucosamine-6-sulfotransferase, useful for treatment,
XX monitoring and diagnosis of macular corneal dystrophy
XX
XX Example 2; Fig 3C; 69pp; English.
XX
XX The present invention describes human corneal
XX N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
XX sulfation of keratan sulfate (KS). Also described is a method for

CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human high endothelial cell
 CC N-acetylglucosamine-6-sulfotransferase (GlcNAc6ST), which is given in
 CC the exemplification of the present invention.

CC Sequence 171 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 23; Length 171;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 129 VRYEDL 134

RESULT 3
 ID ABB81562 standard; Protein; 179 AA.
 AC ABB81562;

DT 05-SEP-2002 (first entry)

DE Human keratan sulfate Gal-6-sulfotransferase portion SEQ ID NO:10.

KW Human: N-acetylglucosamine-6-sulfotransferase; enzyme; GlcNAc6ST;
 KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KW ophthalmological.

OS Homo sapiens.

PN US2002061562-A1.

PD 23-MAY-2002.

PE 09-AUG-2001; 2001US-0927602.

PR 11-AUG-2000; 2000CS-325773P.

PA (FUKU/) FUKUDA M N.
 (AKAM/) AKAMA T O.

PI Fukuda MN, Akama TO;

DR WPI; 2002-507643/54.

XX New nucleic acid encoding corneal

PT N-acetylglucosamine-6-sulfotransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -

PS Example 2; Fig 3C; 69pp; English.

CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfotransferase (I), which is able to catalyze
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratectomy. The present
 CC sequence represents a portion of the human keratan sulfate
 CC Gal-6-sulfotransferase, which is given in the exemplification
 CC of the present invention.

XX Sequence 179 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 23; Length 179;
 Best Local Similarity 100.0%; Pred. No. 52;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 137 VRYEDL 142

RESULT 4
 ID ABB68582 standard; Protein; 183 AA.
 AC ABB68582;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 32538.

KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PE 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;
 WPI; 2001-656860/75.

DR N-PSDB; ABL12685.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 32538; 21pp + Sequence Listing; English.

CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL161/6-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 183 AA;

Query Match Best Local Similarity 100.0%; Score 31; DB 22; Length 183;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
 |||||
 Db 75 VRYEDL 80

RESULT 5
 ID AA011274 standard; Protein; 380 AA.
 AC AA011274;

XX

DT 12-MAR-2002 (first entry)
 XX
 DE Human L-selectin sulfotransferase-2 (LSSR-2) protein.
 XX
 KW Human; beta1,3GNT; beta1,3-N-acetylglucosaminyltransferase; MECA-79;
 KW L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;
 KW ulcerative colitis; inflammatory skin disorder; psoriasis; Lichen planus;
 KW allergic contact dermatitis; lymphoma; chronic pneumonia; LSSR-2;
 KW delayed-type hypersensitivity reaction; hyperplastic thymus; antitumor;
 KW antinflammatory; antipsoriatic; antidiabetic; dermatological;
 KW antiallergic.
 XX
 OS Homo sapiens.
 XX
 PN WO200185177-A1.
 XX
 PD 15-NOV-2001.
 XX
 PF 10-MAY-2001; 2001WO-US15452.
 XX
 PR 11-MAY-2000; 2000US-0569320.
 XX
 PA (BURN-) BURNHAM INST.
 XX
 PI Fukuda M, Yeh J, Hiraoka N;
 XX
 DR WPI; 2002-075226/10.
 XX
 DR N-PSDB; AAS16947.
 XX
 PT New enzyme, useful for modifying acceptor molecule, comprises an
 PT isolated L-selectin sulfotransferase-2 that directs expression of
 PT L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or
 PT intestinal G10NC 6-sulfotransferase
 XX
 PS Claim 21; Fig 4; 98pp; English.
 XX
 CC The present invention provides a method of modifying an acceptor molecule
 CC by contacting the acceptor with an isolated
 CC beta1,3-N-acetylglucosaminyltransferase (beta1,3GNT) or an active
 CC fragment, where beta1,3GNT directs expression of a MECA-79 antigen. The
 CC invention also provides a method of treating or preventing an
 CC L-selectin-mediated condition by reducing the expression or activity of a
 CC beta1,3GNT that directs expression of a MECA-79 antigen. This can be done
 CC by administering to the subject an oligosaccharide L-selectin antagonist
 CC that inhibits binding of L-selectin to a MECA-79 antigen, for example by
 CC administering antibody material that specifically binds beta1,3GNT,
 CC and/or a beta1,3GNT antisense nucleic acid molecule. L-selectin
 CC sulfotransferase-2 (LSSR-2) also directs MECA-79 antigen expression.
 CC Alternatively, the expression or activity of LSSR-2 or its active
 CC fragment can be reduced in combination with reducing the expression or
 CC activity of beta1,3GNT. The method is useful for treating L-selectin
 CC mediated conditions such as Crohn's disease and ulcerative colitis,
 CC inflammatory disorders of the skin such as allergic contact dermatitis,
 CC psoriasis and Lichen planus, lymphomas, chronic pneumonia, delayed-type
 CC hypersensitivity reactions, diabetes and hyperplastic thymus. This
 CC sequence represents human LSSR-2.
 XX
 SQ Sequence 380 AA;
 XX
 QY Query Match 100.0%; Score 31; DB 23; Length 380;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VREDEL 6
 XX
 DB 268 VREDEL 273
 XX
 RESULT 6
 XX AAY39918
 XX ID AAY39918 standard; Protein; 386 AA.
 XX
 AC AAY39918;

XX 08-DEC-1999 (first entry)
 DT
 XX
 DE Human glycosyl sulfotransferase-3 protein sequence.
 XX
 KW Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
 KW selectin binding interaction; inflammation; lymphocyte homing; human;
 KW secondary lymph organ.
 XX
 OS Homo sapiens.
 XX
 PN WO9949018-A1.
 XX
 PD 30-SEP-1999.
 XX
 PE 26-FEB-1999; 99WO-US04316.
 XX
 PR 20-MAR-1998; 98US-0045284.
 XX
 PR 12-NOV-1998; 98US-0190911.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI (SYNT) SYNTX USA INC.
 XX
 PI Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
 XX
 DR WPI; 1999-580442/49.
 XX
 DR N-PSDB; AAZ20792.
 XX
 PT Human and murine glycosyl sulfotransferase 3 and related
 PT polynucleotides
 XX
 PS Claim 2; Fig 1; 59pp; English.
 XX
 CC This sequence is the human glycosyl sulfotransferase-3 (GST-3) of
 CC the invention. The nucleic acid sequences, probes and primers derived
 CC from these, proteins and antibodies are useful in detecting homologues.
 CC The sequences, antibodies and methods are useful in the diagnosis and
 CC treatment of diseases associated with selectin binding interactions;
 CC including conditions associated with or resulting from the homing of
 CC leukocytes to sites of inflammation and the normal homing of lymphocytes
 CC to secondary lymph organs.
 XX
 SQ Sequence 386 AA;
 XX
 QY Query Match 100.0%; Score 31; DB 20; Length 386;
 XX Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 DB 1 VREDEL 6
 XX
 DB 274 VREDEL 279
 XX
 RESULT 7
 XX AAY79219
 XX ID AAY79219 standard; Protein; 386 AA.
 XX
 AC AAY79219;
 XX
 DT 19-JUN-2000 (first entry)
 XX
 DE Human transferase TRNSFS-11.
 XX
 KW Transferase; TRNSFS-11; human; antitumor; cell proliferation;
 KW gastrointestinal disorder; developmental disorder;
 KW genetic disorder; neurological disorder; reproductive disorder;
 KW smooth muscle disorder; immunological disorder; inflammation;
 KW diagnosis; therapy; N-acetylglucosamine 6-O-sulfotransferase.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 121

FT Modified-site /note= "potential O-phosphorylation"
 FT 107
 FT Modified-site /note= "potential O-phosphorylation"
 FT 217
 FT Modified-site /note= "potential O-phosphorylation"
 FT 252
 FT Modified-site /note= "potential O-phosphorylation"
 FT 364
 FT Modified-site /note= "potential O-phosphorylation"
 FT 380
 FT Modified-site /note= "potential O-phosphorylation"
 FT 35
 FT Modified-site /note= "potential O-phosphorylation"
 FT 50
 FT Modified-site /note= "potential O-phosphorylation"
 FT 81
 FT Modified-site /note= "potential O-phosphorylation"
 FT 287
 FT Modified-site /note= "potential O-phosphorylation"
 FT 243
 FT Modified-site /note= "potential O-phosphorylation"
 FT 30
 FT Modified-site /note= "potential N-glycosylation"
 FT 308
 FT Modified-site /note= "potential N-glycosylation"
 FT 329
 FT Modified-site /note= "potential N-glycosylation"
 FT 7...23
 FT Domain /note= "potential N-glycosylation"
 FT /note= "transmembrane domain"
 PN WO200014251-A2.
 XX 16-MAR-2000.
 PD 09-SEP-1999; 99WO-US20989.
 XX 10-SEP-1998; 98US-0150657.
 PR 04-NOV-1998; 98US-0186779.
 PR 11-MAY-1999; 98US-0133642.
 XX (INCY-) INCYTE PHARM INC.
 PA Tang YT, Corley NC, Guegler KJ, Baughn MR, Lal P, Yue H;
 PI Hillman JL, Azimzai Y;
 PI Hillman JL, Azimzai Y;
 XX WPI: 2000-256996/22.
 DR N-PSDB; AAZ94211.
 XX Human transferase proteins useful for preventing, diagnosing and
 PT treating cancers and developmental, gastrointestinal, genetic,
 PT immunological, neurological, reproductive and smooth muscle disorders -
 XX
 PS Claim 1; Page 90-91; 113pp; English.
 XX The present sequence is that of human transferase TRNSF-11, 1 of
 CC 15 claimed novel human transferase proteins of the invention (see
 CC AAY79209-23). The sequence was deduced from a cDNA clone (see
 CC AAZ94211) isolated from a gallbladder library. It shows homology to
 CC mouse N-acetylglucosamine 6-O-sulfotransferase. TRNSF-11 is
 CC expressed in dermatologic and gastrointestinal tissues,
 CC especially those associated with inflammation and cell
 CC proliferation. The new human transferases and polynucleotides can
 CC be used in the diagnosis, prevention and treatment of cancer,
 CC developmental disorders, gastrointestinal disorders, genetic
 CC disorders, immunological disorders, neurological disorders,
 CC reproductive disorders, and smooth muscle disorders. The
 CC polypeptides can also be used to raise antibodies, and to screen
 CC for agonists and antagonists of transferase activity.
 XX
 SQ Sequence 386 AA;

Query Match 100.0%; Score 31; DB 21; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 DB 274 VRYEDL 279
 RESULT 8
 AAM93309
 ID AAM93309 standard; Protein; 386 AA.
 AC AAM93309;
 XX
 DT 06-NOV-2001 (first entry)
 DE Human polypeptide; SEQ ID NO: 2817.
 KW Human; full length cDNA; cDNA synthesis; oligo-capping.
 XX Homo sapiens.
 OS
 PN EP1130094-A2.
 PD 05-SEP-2001.
 PF 07-JUL-2000; 2000EP-0114089.
 XX
 PR 08-JUL-1999; 98JP-0194486.
 PR 11-JAN-2000; 2000JP-0118774.
 PR 02-MAY-2000; 2000JP-0183765.
 XX
 PA (HELI-) HELIX RES INST.
 PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
 PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;
 XX
 DR WPI: 2001-524255/58.
 DR N-PSDB; AAK944229.
 XX
 PT 830 Primers useful for synthesizing full length cDNA clones and their
 PT use in genetic manipulation -
 XX
 PS Claim 8; SEQ ID NO 2817; 1380pp + sequence listing; English.
 XX
 CC The invention relates to primers for synthesizing full length cDNA
 CC clones. 830 cDNA molecules encoding a human protein have been
 CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
 CC molecules have been determined. Primers for synthesizing the full length
 CC cDNA are useful for clarifying the function of the protein encoded by
 CC the cDNA. The full length clones were obtained by construction of full
 CC length enriched cDNA libraries that were synthesised by the oligo-capping
 CC method. The primers enable the production of the full length cDNA easily
 CC without any special methods. The present sequence is a polypeptide
 CC encoded by a full length human cDNA of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in CD-ROM format directly from EPO.
 XX
 SQ Sequence 386 AA;

Query Match 100.0%; Score 31; DB 22; Length 386;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYEDL 6
 DB 274 VRYEDL 279
 RESULT 9
 AAY72638
 ID AAY72638 standard; Protein; 395 AA.
 AC AAY72638;
 XX

```

XX 02-MAY-2001 (first entry)
DT
XX Mouse glycosyl sulfotransferase-4 (GST-4).
DE
XX Mouse; glycosyl sulfotransferase-4; GST-4; immunosuppressive;
XX therapy; selectin binding inhibitor; gene therapy; inflammation;
XX systemic lupus erythematosus; SLE; rheumatoid arthritis; diabetes;
XX polyarthritis nodosa; polymyositis; systemic sclerosis; dermatitis;
XX glomerulonephritis; myasthenia gravis; Sjogren's syndrome; adrenalitis;
XX Hashimoto's disease; Grave's disease; hypoparathyroidism; anaemia;
XX demyelinating disease; cirrhosis; ulcerative colitis; allergic rhinitis;
XX myocarditis; adult respiratory distress syndrome; eczema; psoriasis;
XX asthma; hypersensitivity; rheumatic fever; tissue rejection;
XX chromosome 8E1.
XX
XX Mus musculus.
XX
XX WO200106015-A1.
XX
XX 25-JAN-2001.
XX
XX 19-JUL-2000; 2000WO-US19741.
XX
XX 20-JUL-1999; 98US-0144694.
XX 13-JUL-2000; 2000US-0393828.
XX
XX (REGC ) UNIV CALIFORNIA.
XX
XX Rosen SD, Lee JK, Hemmerich S;
XX
XX WPI: 2001-138472/14.
XX DR N-PSDB; AAD02696.
XX
XX New glycosyl sulfotransferases (GST)-4alpha, GST-4beta and GST-6 for
XX diagnostic and therapeutic agent screening applications -
XX
XX Claim 3; Fig 2; 128pp; English.
XX
XX The present sequence is mouse glycosyl sulfotransferase-4 (GST-4). GST-4
XX gene is found on chromosome 8E1.
XX
XX GST is a type 2 membrane protein useful for inhibiting a binding event
XX between a selectin and a selectin ligand, which comprises contacting the
XX selectin with a non-sulphated selectin ligand. GST and a small molecular
XX agent that inhibits the sulphation activity of GST. GST is also useful
XX in inhibiting a selectin mediated binding event. GST is useful in gene
XX therapy to treat disorders such as acute or chronic inflammation,
XX systemic lupus erythematosus (SLE), rheumatoid arthritis, polyarteritis
XX nodosa, polymyositis, dermatomyositis, systemic sclerosis, diabetes,
XX glomerulonephritis, myasthenia gravis, Sjogren's syndrome, Hashimoto's
XX disease, Grave's disease, adrenalitis, hypoparathyroidism, pernicious
XX anemia, demyelinating diseases, cirrhosis, ulcerative colitis,
XX dermatitis, myocarditis, regional enteritis, adult respiratory distress
XX syndrome, infantile eczema, psoriasis lichen planus, allergic rhinitis,
XX bronchial asthma, hypersensitivity, rheumatic fever and tissue rejection
XX during transplantation.
XX
XX SO Sequence 395 AA;

```

Query Match 100.0%; Score 31; DB 22; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VRYEDL 6
   |||||
Db 271 VRYEDL 276

```

RESULT 10
 AAU11275
 ID AAU11275 standard; Protein; 395 AA.
 XX
 AC AAU11275;

```

XX 12-MAR-2002 (first entry)
DT
XX Murine intestinal-GlcNAc 6-sulfotransferase (I-GlcNAc6ST) protein.
DE
XX Mouse; beta1,3GNT; beta1,3-N-acetylglucosaminyltransferase; MECA-79;
XX L-selectin; L-selectin sulfotransferase-2; Crohn's disease; diabetes;
XX ulcerative colitis; inflammatory skin disorder; psoriasis; lichen planus;
XX allergic contact dermatitis; lymphoma; chronic pneumonia; LSS1-2;
XX delayed-type hypersensitivity reaction; hyperplastic thymus; antileuk;
XX anti-inflammatory; antipsoriatic; antididiabetic; dermatological;
XX antiallergic; intestinal-GlcNAc 6-sulfotransferase; I-GlcNAc6ST.
XX
XX Mus musculus.
XX
XX WO200185177-A1.
XX
XX 15-NOV-2001.
XX
XX 10-MAY-2001; 2001WO-US15452.
XX
XX 11-MAY-2000; 2000US-0569320.
XX
XX (BURN-) BURNHAM INST.
XX
XX Fukuda M, Yeh J, Hiraoka N;
XX
XX WPI: 2002-075226/10.
XX DR N-PSDB; AAS16948.
XX
XX New enzyme, useful for modifying acceptor molecule, comprises an
XX isolated L-selectin sulfotransferase-2 that directs expression of
XX L-selectin ligand antigen, MECA-79 in Chinese hamster ovary cells, or
XX intestinal GlcNAc 6-sulfotransferase -
XX
XX Claim 28; Fig 10; 98pp; English.
XX
XX The present invention provides a method of modifying an acceptor molecule
XX by contacting the acceptor with an isolated
XX beta1,3-N-acetylglucosaminyltransferase (beta1,3GNT) or an active
XX fragment, where beta1,3GNT directs expression of a MECA-79 antigen. The
XX invention also provides a method of treating or preventing an
XX L-selectin-mediated condition by reducing the expression or activity of a
XX beta1,3GNT that directs expression of a MECA-79 antigen. This can be done
XX by administering to the subject an oligosaccharide L-selectin antagonist
XX that inhibits binding of L-selectin to a MECA-79 antigen, for example by
XX administering antibody material that specifically binds beta1,3GNT,
XX and/or a beta1,3GNT antisense nucleic acid molecule. L-selectin
XX sulfotransferase-2 (LSS1-2) also directs MECA-79 antigen expression.
XX Alternatively, the expression or activity of LSS1-2 or its active
XX fragment can be reduced in combination with reducing the expression or
XX activity of beta1,3GNT. The method is useful for treating L-selectin
XX mediated conditions such as Crohn's disease and ulcerative colitis,
XX inflammatory disorders of the skin such as allergic contact dermatitis,
XX psoriasis and lichen planus, lymphomas, chronic pneumonia, delayed-type
XX hypersensitivity reactions, diabetes and hyperplastic thymus. This
XX sequence represents mouse I-GlcNAc6ST.
XX
XX SO Sequence 395 AA;

```

Query Match 100.0%; Score 31; DB 23; Length 395;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 VRYEDL 6
   |||||
Db 271 VRYEDL 276

```

RESULT 11
 AAW61100
 ID AAW61100 standard; Protein; 411 AA.
 XX

AC AAW61100;
 XX
 DT 14-SEP-1998 (first entry)
 XX
 DE Keratan sulphate 6-sulphotransferase.
 XX
 KM KSGal6ST; keratan sulphate 6-sulphotransferase;
 KM chick chondroitin 6-sulphotransferase; CGST; phage lambda; hybridization;
 KM expression vector; COS-7 cells; Bluescript plasmid; galactose;
 KM keratan sulphate; chondroitin.
 XX
 OS Homo sapiens.
 XX
 PN EP845533-A2.
 XX
 PD 03-JUN-1998.
 XX
 PF 27-NOV-1997; 97EP-0309564.
 XX
 PR 29-NOV-1996; 96JP-0320535.
 XX
 PA (SECK) SEIKAGAKU KOGYO CO LTD.
 XX
 PI Fukuta M, Habuchi O;
 XX
 DR WPI; 1998-288750/26.
 DR N-PSDB; AAV36418.
 XX
 PT Keratan sulphate 6-sulpho-transferase polypeptide - transfers
 PT sulphate from sulphate donor to galactose 6-hydroxy group etc.
 XX
 PS Claim 4; Page 15-16; 21pp; English.
 XX
 CC This sequence produces the protein KSGal6ST (Keratan sulphate
 CC 6-sulphotransferase), which has a molecular weight of about 46700. The
 CC cDNA sequence of this protein was obtained by radiolabelling the cDNA of
 CC chick chondroitin 6-sulphotransferase (CGST) and using this as a probe in
 CC a random oligonucleotide-primed labelling method. Human foetal brain
 CC cDNA was inserted into a phage lambda g11 cloning vector whereby the
 CC clones containing the KSGal6ST were obtained by hybridization using the
 CC prepared probe. The positive clones were subcloned into a recombinant
 CC expression vector and used to transform COS-7 cells, from which cells,
 CC expressing KSGal6ST can be selected. The phage cDNA inserts were
 CC isolated and subcloned into a Bluescript plasmid. Deletion clones were
 CC then prepared from which both strands were sequenced by the Sanger
 CC method. The KSGal6ST of the invention transfers the sulphate from a
 CC sulphate donor to galactose 6-OH groups in keratan sulphate, but does not
 CC transfer sulphate to chondroitin, chondroitin sulphate A or C, dermatan
 CC sulphate or CDSNS heparin.
 XX
 SQ Sequence 411 AA;
 XX
 Query Match 100.0%; Score 31; DB 19; Length 411;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 300 VRYEDL 305
 XX
 RESULT 12
 ABB81557
 ID ABB81557 standard; Protein; 418 AA.
 XX
 AC ABB81557;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Mouse intestinal N-acetylglucosamine-6-sulfoyltransferase SEQ ID NO:5.
 XX
 KM Human; N-acetylglucosamine-6-sulfoyltransferase; enzyme; GlcNAc6ST;
 KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 XX

KM ophthalmological.
 XX
 OS Mus musculus.
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX
 PF 09-AUG-2001; 2001US-0927602.
 XX
 PR 11-AUG-2000; 2000US-325773P.
 XX
 PA (FUKU/) FUKUDA M.N.
 PA (AKAM/) AKAMA T.O.
 XX
 PI Fukuda MN, Akama TO;
 XX
 DR WPI; 2002-507643/54.
 XX
 PT New nucleic acid encoding corneal
 PT N-acetylglucosamine-6-sulfoyltransferase, useful for treatment,
 PT monitoring and diagnosis of macular corneal dystrophy -
 XX
 PS Example 5; Page 24-25; 69pp; English.
 XX
 CC The present invention describes human corneal
 CC N-acetylglucosamine-6-sulfoyltransferase (I), which is able to catalyse
 CC sulfation of keratan sulfate (KS). Also described is a method for
 CC monitoring the effect of treatments for macular corneal dystrophy (MCD),
 CC and detecting susceptibility to MCD. (I) is located to chromosome 16q22,
 CC and has ophthalmological activity. (I) can be used to treat or prevent
 CC macular corneal dystrophy types I or II. (I) makes possible treatment
 CC of MCD without requiring keratoplasty or keratotomy. The present
 CC sequence represents mouse intestinal N-acetylglucosamine-6-
 CC sulfoyltransferase, which is given in comparison with (I) in the
 CC exemplification of the present invention.
 XX
 SQ Sequence 418 AA;
 XX
 Query Match 100.0%; Score 31; DB 23; Length 418;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 294 VRYEDL 299
 XX
 RESULT 13
 AAY31656
 ID AAY31656 standard; Protein; 483 AA.
 XX
 AC AAY31656;
 XX
 DT 09-NOV-1999 (first entry)
 XX
 DE Mouse N-acetylglucosamine-6-O-sulfoyltransferase.
 XX
 KM N-acetylglucosamine-6-O-sulfoyltransferase; mouse; GlyCAM-1;
 KM L-selectin ligand.
 XX
 OS Mus musculus.
 XX
 PN EP943668-A2.
 XX
 PD 22-SEP-1999.
 XX
 PF 04-MAR-1999; 99BP-0301530.
 XX
 PR 24-JUN-1998; 98JP-0177844.
 PR 05-MAR-1996; 98JP-0054007.
 XX
 PA (SECK) SEIKAGAKU CORP.
 XX

```
XX Habuchi O, Kadomatsu K, Kannagi R, Muramatsu H;
PI Muramatsu T, Uchimura K;
XX WPI: 1999-520337/44.
DR N-PSDB: AAX87820.
XX
PT New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PI for synthesis of sugar chains, e.g. GLYCAM-1
XX
PS Claim 1; Page 24-25; 41pp; English.
XX
CC The present sequence represents mouse N-acetylglucosamine-6-O-
CC sulfotransferase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC beta 1-3 = beta 1-3 glycosidic linkage, and beta 1-4 = beta 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GlyCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87820) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GlyCAM-1 using transformants which harbour
CC the DNA.
CC
SQ Sequence 483 AA;
XX
Query Match 100.0%; Score 31; DB 20; Length 483;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VREYDL 6
DB 376 VREYDL 381
XX
RESULT 14
AAY31657
ID AAY31657 standard; Protein; 484 AA.
XX
AC AAY31657;
XX
DT 09-NOV-1999 (first entry)
XX
DE Human N-acetylglucosamine-6-O-sulfotransferase.
XX
KM N-acetylglucosamine-6-O-sulfotransferase; human; GlyCAM-1;
XX L-selectin ligand.
XX
OS Homo sapiens.
XX
PN EP943668-A2.
XX
PD 22-SEP-1999.
XX
PF 04-MAR-1999; 99EP-0301530.
XX
PR 24-JUN-1998; 98JP-0177844.
XX
PR 05-MAR-1998; 98JP-0054007.
XX
PA (SEGG) SEIKAGAKU CORP.
XX
PI Habuchi O, Kadomatsu K, Kannagi R, Muramatsu H;
PI Muramatsu T, Uchimura K;
XX
XX WPI: 1999-520337/44.
XX
XX N-PSDB: AAX87821.
XX
PT New polypeptides of N-acetylglucosamine-6-O-sulfotransferase, useful
PI for synthesis of sugar chains, e.g. GlyCAM-1
```

```
XX Claim 2; Page 28-30; 41pp; English.
PS
XX
CC The present sequence represents human N-acetylglucosamine-6-O-
CC sulfotransferase, an enzyme capable of transferring a sulfate group
CC from a sulfate group donor to a hydroxyl group at the 6 position of
CC an N-acetylglucosamine residue located at the non-reducing end of
CC an oligosaccharide of formula GlcNAc $\beta$ 1-3Gal $\beta$ 1-4GlcNAc, where
CC GlcNAc = N-acetylglucosamine residue, Gal = galactose residue,
CC beta 1-3 = beta 1-3 glycosidic linkage, and beta 1-4 = beta 1-4
CC glycosidic linkage. The enzyme is useful for the synthesis of
CC sugar chains such as GlyCAM-1, a ligand of L-selectin that is
CC involved in homing of lymphocytes and rolling of leukocytes
CC occurring at the early stage of inflammation. DNA encoding the
CC enzyme (see AAX87821) is expected to be used for the large-scale
CC production of N-acetylglucosamine-6-O-sulfotransferase, or
CC artificial synthesis of GlyCAM-1 using transformants which harbour
CC the DNA.
CC
SQ Sequence 484 AA;
XX
Query Match 100.0%; Score 31; DB 20; Length 484;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 VREYDL 6
DB 377 VREYDL 382
XX
RESULT 15
AAB95367
ID AAB95367 standard; Protein; 530 AA.
XX
AC AAB95367;
XX
DT 26-JUN-2001 (first entry)
XX
DE Human protein sequence SEQ ID NO:17679.
XX
KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX
OS Homo sapiens.
XX
PN EP1074617-A2.
XX
PD 07-FEB-2001.
XX
PF 28-JUL-2000; 2000EP-0116126.
XX
PR 29-JUL-1999; 99JP-0248036.
XX
PR 27-AUG-1999; 99JP-0300253.
XX
PR 11-JAN-2000; 2000JP-0118776.
XX
PR 02-MAY-2000; 2000JP-0183767.
XX
PR 09-JUN-2000; 2000JP-0241899.
XX
PA (HELI-) HELIX RES INST.
XX
PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
PT Primer sets for synthesizing polynucleotides, particularly the 5602
PT full-length cDNAs defined in the specification, and for the detection
PT and/or diagnosis of the abnormality of the proteins encoded by the
PT full-length cDNAs -
XX
XX Claim 8; SEQ ID 17679; 2537pp + CD ROM; English.
XX
XX The present invention describes primer sets for synthesizing 5602
XX full-length cDNAs defined in the specification. Where a primer set
XX comprises: (a) an oligo-dT primer and an oligonucleotide complementary
```

CC to the complementary strand of a polynucleotide which comprises one of
CC the 5602 nucleotide sequences defined in the specification, where the
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
CC of an oligonucleotide comprising a sequence complementary to the
CC complementary strand of a polynucleotide which comprises a 5'-end
CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end sequence, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in
CC the specification. The primer sets can be used in antisense therapy and
CC in gene therapy. The primers are useful for synthesising polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
CC represent oligonucleotides, all of which are used in the exemplification
CC of the present invention.

XX SQ Sequence 530 AA;

Query Match 100.0%; Score 31; DB 22; Length 530;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

Db 423 VRYEDL 428

RESULT 16

AAU69414

ID AAU69414 standard; Protein; 531 AA.

XX AC AAU69414;

XX DT 30-JAN-2002 (first entry)

XX DE Lung small cell carcinoma antigen #8.

XX KM Human; cytostatic; antitumour; lung small cell cancer antigen;

XX KW tumour; lung cancer.

XX OS Homo sapiens.

XX PN WO200177168-A2.

XX PD 18-OCT-2001.

XX PF 11-APR-2001; 2001WO-US11859.

XX PR 11-APR-2000; 2000US-196780P.

XX PR 21-JUN-2000; 2000US-21361P.

XX PR 01-SEP-2000; 2000US-229763P.

XX PR 05-SEP-2000; 2000US-230629P.

XX PR 14-SEP-2000; 2000US-232656P.

XX PR 19-DEC-2000; 2000US-257037P.

XX PR 08-JAN-2001; 2001US-260796P.

XX PA (CORI-) CORIXA CORP.

XX PI Lodes MJ, Wang T, Mohamath R, Indirias CY;

XX PS Claim 1; Page 219-220; 295pp; English.

XX CC The invention relates to novel isolated lung small cell cancer antigen
CC polynucleotides (I) and polypeptides (II) used in a method of detecting

CC cancer in a patient. The method is optionally performed by
CC utilising oligonucleotides (III), where the biological sample
CC from the patient is contacted with (III), detecting the amount of
CC polynucleotide hybridised to (III) in the sample and comparing the
CC amount of polynucleotide to a predetermined cut-off value and thereby
CC determining cancer in a patient. (I), (II) or antigen-presenting cells
CC expressing (II) is useful for stimulating and/or expanding T cells
CC specific for a tumour protein. The method comprises contacting T cells
CC with one of the components under conditions to permit the stimulation
CC and/or expansion of the cells. A composition comprising (I) is useful for
CC stimulating an immune response in a patient and for inhibiting the
CC development of a cancer especially lung cancer in a patient. An
CC isolated T cell population is useful for removing tumour cells from the
CC biological sample and for inhibiting the development of cancer in a
CC patient. AAU69407-AAU69431 represent novel human lung small cell
CC cancer antigen amino acid sequences of the invention.

XX SQ Sequence 531 AA;

Query Match 100.0%; Score 31; DB 23; Length 531;

Best Local Similarity 100.0%; Pred. No. 1.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

Db 424 VRYEDL 429

RESULT 17

ABB67037

ID ABB67037 standard; Protein; 649 AA.

XX AC ABB67037;

XX DT 26-MAR-2002 (first entry)

XX DE Drosophila melanogaster polypeptide SEQ ID NO 27903.

XX KW Drosophila; developmental biology; cell signalling; insecticide;

XX KM pharmaceutical.

XX OS Drosophila melanogaster.

XX PN WO200171042-A2.

XX PD 27-SEP-2001.

XX PF 23-MAR-2001; 2001WO-US09231.

XX PR 23-MAR-2000; 2000US-191637P.

XX PR 11-JUL-2000; 2000US-0614150.

XX PA (PEKE) PE CORP NY.

XX PI Venter JC, Adams M, Li PMD, Myers EW;

XX PR N-PSDB; ABL11140.

XX DR WPI: 2001-656860/75.

XX PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -

XX PS Disclosure; SEQ ID NO 27903; 21pp + Sequence Listing; English.

XX CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL101840-ABL16175) and the encoded proteins
XX (ABB57737-ABB72072).

CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

CC Sequence 649 AA;

Query Match 100.0%; Score 31; DB 22; Length 649;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 :|||||
 DB 332 VREEDL 337

RESULT 18
 ABB64512
 ID ABB64512 standard; Protein: 363 AA.

XX ABB64512;

XX 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 20328.

KW Drosophila: developmental biology; cell signalling; insecticide;
 KW pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL08615.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -

PS Disclosure; SEQ ID NO 20328; 21pp + Sequence listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB57737-AB872072).
 CC (AB57737-AB872072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 363 AA;

Query Match 96.8%; Score 30; DB 22; Length 363;
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 :|||||
 DB 258 IREEDL 263

RESULT 19
 ABB09452
 ID ABB09452 standard; Protein: 562 AA.

XX ABB09452;

XX 01-JUL-2002 (first entry)

DE Glucose-3-dehydrogenase amino acid sequence.

KW Glucose-3-dehydrogenase; G3DH; enzyme.

OS Halomonas sp. alpha-15.

PN JP2002017372-A.

PD 22-JAN-2002.

PF 30-JUN-2000; 2000JP-0237709.

PR 30-JUN-2000; 2000JP-0237709.

PA (HAYADA) HAYADA K.

DR WPI; 2002-221711/28.

DR N-PSDB; ABL52744.

PT Glucose-3-dehydrogenase and its preparation -

PS Claim 1a; Page 7-9; 16pp; Japanese.

XX The invention relates to Glucose-3-dehydrogenase (G3DH) enzyme. The
 CC methods of the invention are used for the preparation of G3DH. The
 CC current sequence represents the glucose-3-dehydrogenase amino acid
 CC sequence.

XX Sequence 562 AA;

Query Match 96.8%; Score 30; DB 23; Length 562;
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 :|||||
 DB 142 IREEDL 147

RESULT 20
 ABB92113
 ID ABB92113 standard; Protein: 405 AA.

XX ABB92113;

XX 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 1324.

KW Herbicidally active polypeptide; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;
 XX WPI; 2002-269010/31.
 DR
 XX
 PT Identifying plant target proteins for herbicidally active compounds,
 PT comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 PS
 XX Claim 5; SEQ ID NO 1324; 261pp + Sequence Listing; English.
 CC
 CC The invention relates to identifying target proteins
 CC (AB930790-AB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.
 CC
 SQ Sequence 405 AA;
 Query Match 93.5%; Score 29; DB 23; Length 405;
 Best Local Similarity 83.3%; Pred. No. 3.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 335 VRYEDL 340
 RESULT 21
 ABP32235
 ID ABP32235 standard; Protein; 139 AA.
 XX
 XX ABP32235;
 AC
 DT 09-JUL-2002 (first entry)
 XX
 DE Human isomerase-like ORF1208 protein, SEQ ID NO:2416.
 XX
 KM Human; ORF; open reading frame; ORFX; drug screening; diagnosis;
 KM disease monitoring; cytokine; cell proliferation; cell differentiation;
 KM immune modulation; haematopoiesis regulation; tissue growth;
 KM angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KM thrombolytic; tumour inhibition; bodily characteristics; fertility;
 KM behaviour; cancer; proliferative disorder; neurological disorder;
 KM cardiovascular disease; immune system disorder; organ transplantation;
 KM tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KM hypothyroidism; cholesterol ester storage disease; infection; vulnery;
 KM vasotropic; antiproliferative; antidiabetic; cyostatic; nootropic;
 KM neuroprotective; antithrombotic; anticoagulant; thrombolytic;
 KM cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KM dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Leach MD, Shinkets RA;
 XX
 DR WPI; 2002-106200/14.
 DR N-PSDB; ABN76261.
 XX

PT Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and disorders related to organ
 PT transplantation -
 PS
 XX Claim 10; Page 866; 2508pp; English.
 CC
 CC Sequences ABP31028-ABP35561 represent 4534 novel human proteins
 CC designated ORF (open reading frame) 1-4534, and sequences ABN75054-
 CC ABN79587 represent cDNAs encoding them. The invention also encompasses
 CC polypeptides at least 80% identical to the ORF1-ORF4534 (collectively
 CC referred to as ORFX) proteins, polynucleotides at least 85% identical to
 CC the ORFX nucleic acid sequences, vectors and host cells comprising ORFX
 CC polynucleotides, the recombinant production of ORFX proteins, antibodies
 CC specific for ORFX proteins, methods of detecting ORFX polynucleotides and
 CC polypeptides, methods of screening for modulators of ORFX expression or
 CC activity, and methods of screening individuals for a predisposition to an
 CC ORFX-associated disorder. The ORFX proteins of the invention have a wide
 CC range of biological activities, such as cytokine, cell proliferation,
 CC cell differentiation, immune modulation, haematopoiesis regulation,
 CC tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic activity, thrombolytic activity,
 CC receptor/ligand, antinflammatory activity, tumour inhibition activity,
 CC and antineoplastic activity, and may also be involved in the determination
 CC of bodily characteristics, fertility and behaviour. ORFX proteins,
 CC nucleic acids and antibodies may be used in the treatment of cancers,
 CC other proliferative disorders such as psoriasis and benign tumours,
 CC neurological disorders such as epilepsy and Alzheimer's disease,
 CC cardiovascular diseases, immune system disorders, disorders related to
 CC organ transplantation, disorders of tissue growth and regeneration,
 CC diseases such as diabetes mellitus, hypothyroidism, and cholesterol ester
 CC storage disease, and infectious diseases caused by viral, bacterial,
 CC fungal and other pathogens. ORFX nucleic acids may also be used as a
 CC source of primers and probes, in the detection of ORFX genomic sequences
 CC or transcripts, in the identification and cloning of homologous
 CC sequences, in genetic diagnosis, and in forensic biology. The ORFX
 CC nucleic acids may additionally be used to produce transgenic animals
 CC which may be useful for studying the function and/or activity of ORFX
 CC protein, and in drug screening. The ORFX proteins may also be used as
 CC immunogens to generate specific antibodies, which are useful in the
 CC diagnosis, treatment and monitoring of ORFX-associated diseases.
 CC
 SQ Sequence 139 AA;
 Query Match 90.3%; Score 28; DB 23; Length 139;
 Best Local Similarity 83.3%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 123 LRYEDL 128
 RESULT 22
 ABB81563
 ID ABB81563 standard; Protein; 174 AA.
 XX
 AC ABB81563;
 XX
 DT 05-SEP-2002 (first entry)
 XX
 DE Human chondroitin-6-sulfotransferase portion SEQ ID NO:11.
 XX
 KM Human; N-acetylgalucosamine-6-sulfotransferase; enzyme; GICNAC65T;
 KM corneal; sulfation; keratan sulfate; macular corneal dystrophy; MCD;
 KM ophthalmological.
 XX
 OS Homo sapiens.
 XX
 PN US2002061562-A1.
 XX
 PD 23-MAY-2002.
 XX

CC frequent use of broad spectrum antibiotics.
 XX
 SQ Sequence 254 AA;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 22; Length 254;
 Pred. No. 3.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Caps 0;

OY 1 VRYEDL 6
 |||||
 Db 143 VRYQDL 148

RESULT 25
 AAR83040
 ID AAR83040 standard; peptide; 306 AA.
 XX
 AC AAR83040;

DT 03-APR-1996 (first entry)

XX Capsular polysaccharide of Streptococcus pneumoniae (CpsU).
 XX
 DE Capsular polysaccharide; cps; peptide; flanking region; detection;
 KM serotype; diagnosis; prevention; Streptococcus pneumoniae.
 XX

OS Streptococcus pneumoniae.

XX WO9531548-A1.

XX 23-NOV-1995.

XX 16-MAY-1995; 95WO-US06119.

XX 16-MAY-1994; 94US-0243546.

XX (UABR-) UAB RES FOUND.

XX Dillard J, Yother J;

XX WPI; 1996-010934/01.

XX N-PSDB; AA105848.

XX New Streptococcus pneumoniae capsular polysaccharide genes - used
 PT for detection, serotyping and for diagnosis and prevention of S.
 PT pneumoniae infection

PS Disclosure; Page 178-180; 226pp; English.

XX Sequences encoding the 5' flanking region of the capsular
 CC polysaccharide gene (cps) of Streptococcus pneumoniae and which are
 CC of sufficient length to allow hybridisation under standard
 CC hybridisation conditions to a S.pneumoniae cps gene flanking region
 CC may also be used in methods to detect and serotype S.pneumoniae. They
 CC may also be used for the diagnosis and prevention of S. pneumoniae
 CC infection.

SQ Sequence 306 AA;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 17; Length 306;
 Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 165 VRYEDV 170

RESULT 26
 AAY33895
 ID AAY33895 standard; Protein; 306 AA.
 XX
 AC AAY33895;

XX 04-JAN-2000 (first entry)
 XX
 DE Amino acid sequence of Cps3U.

XX Infection; bacterial capsule; serotype specific; virulence;
 KM diagnosis; capsule synthesis; cassette model; binary model;
 KW cassette; in; flanking region; non-serotype specific.

OS Streptococcus pneumoniae.

XX US5948900-A.

XX 07-SEP-1999.

XX 02-JUN-1997; 97US-0867030.

XX 02-JUN-1997; 97US-0867030.

XX 16-MAY-1994; 94US-0243546.

XX (UABR-) UAB RES FOUND.

XX Dillard J, Yother J;

XX WPI; 1999-517979/43.

XX N-PSDB; AA211802.

XX Non-serotype-specific flanking nucleotide sequences derived from
 PT Streptococcus pneumoniae capsular polysaccharide genes, useful as
 PT hybridization probes for identifying serotype-specific capsular
 PT polysaccharide genes

PS Example 10; Fig 6g; 140pp; English.

XX This is the amino acid sequence of the capsule synthesis protein,
 CC Cps3U.

CC Serotype-specific cps genes encode the various enzymatic
 CC functions of capsule synthesis and hence determine the particular

CC structure of the capsule polysaccharides produced. Therefore cps genes
 CC determine the serotype of the virus.

CC The non-serotype specific flanking nucleotide sequences (AA211798,
 CC AA211799, AA211800, AA211801 and AA211802) may be used in hybridization

CC assays to identify the location of DNA flanking serotype-specific cps
 CC genes in any strain of S. pneumoniae virus.

CC Additionally, the flanking regions are involved in recombination
 CC and integration of the type specific cps genes during virus

CC replication. Therefore, when a selected gene (such as the
 CC serotype-specific cps genes) is integrated between the nucleic acids

CC of the flanking region, the resulting construct can be stably integrated
 CC into an S.pneumoniae host.

CC This allows the locus of the serotype-specific cps genes (and other
 CC genetic elements), within a virus, to be identified, and therefore,

CC isolated and characterized. The cps gene sequences may then be used as
 CC hybridization probes to identify and quantify specific serotypes of S.

CC pneumoniae or to recombinantly produce the gene products for use as
 CC antigens in immunization protocols.

SQ Sequence 306 AA;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 20; Length 306;
 Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 Db 165 VRYEDV 170

RESULT 27
 ABB91753
 ID ABB91753 standard; Protein; 333 AA.
 XX
 AC ABB91753;

PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147316.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154039.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156558.
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0158655.
PR 07-OCT-1999; 99US-0158023.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161993.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 349;
Best Local Similarity 93.3%; Pred. No. 5e+02; Mismatches 0; Gaps 0;
Matches 5; Conservative 1; Indels 0;

QY 1 VRYEDL 6
Db 244 LRYEDL 249

RESULT 29
AAC46838
ID AAC46838 standard; Protein; 349 AA.
XX
XX AAC46838;
AC
XX 18-OCT-2000 (first entry)
DE
XX Arabidopsis thaliana protein fragment SEQ ID NO: 58967.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
PN EP1033405-A2.
XX
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 29-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132468.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137774.
PR 08-JUN-1999; 99US-0138054.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140685.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150984.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.

PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160880.
PR 22-OCT-1999; 99US-0160881.
PR 22-OCT-1999; 99US-0160981.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 349;
Best Local Similarity 83.3%; Pred. No. 56+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 244 LRYEDL 249

RESULT 30
AAG44206
ID AAG44206 standard; Protein; 354 AA.
XX AAG44206;

AC AAG44206;
XX
XX 18-OCT-2000 (first entry)
DT
XX
XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55344.

XX
XX Protein identification: signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
KM termination sequence.

OS Arabidopsis thaliana.

XX
XX
PN EPI033405-A2.

XX
XX
PD 06-SEP-2000.

XX
XX
PE 25-FEB-2000; 2000EP-0301439.

XX
XX
PF 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131442.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134370.
PR 14-MAY-1999; 99US-0134376.
PR 18-MAY-1999; 99US-0134941.
PR 19-MAY-1999; 99US-0135124.
PR 20-MAY-1999; 99US-0135353.
PR 21-MAY-1999; 99US-0136021.
PR 24-MAY-1999; 99US-0136782.
PR 25-MAY-1999; 99US-0136782.
PR 27-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139859.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158237.
PR 12-OCT-1999; 99US-0158363.
PR 13-OCT-1999; 99US-0158929.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 13-OCT-1999; 99US-0159299.
PR 14-OCT-1999; 99US-0159320.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.

PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 26-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 354;
Best Local Similarity 83.3%; Pred. No. 5e-02; Mismatches 0; Gaps 0;
Matches 5; Conservative 1; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 31
AAG46837 standard; Protein; 354 AA.
XX AAG46837;
AC AAG46837;
XX 18-OCT-2000 (first entry)
DT 18-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 58966.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58966.
XX Protein identification: signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX EPI033405-A2.
PN EPI033405-A2.
XX 06-SEP-2000.
PD 06-SEP-2000.
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-2000; 2000EP-0301439.
XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134321.
PR 14-MAY-1999; 99US-0134770.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139839.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140391.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142877.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145216.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 28-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149175.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158365.
PR 13-OCT-1999; 99US-0158923.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159684.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.

PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161922.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 354;
Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 32
ABR93382
ID ABR93382 standard; Protein: 354 AA.

AC ABR93382;

DT 31-MAY-2002 (first entry)

DE Herbicidally active polypeptide SEQ ID NO 2593.

KW Herbicidal; plant; agriculture; herbicide.

OS Arabidopsis thaliana.

PN WO200210210-A2.

PD 07-FEB-2002.

PF 28-AUG-2001; 2001WO-EP09892.

PR 28-AUG-2001; 2001WO-EP09892.

PA (FARB) BAYER AG.

PI Tietjen K, Weidler M;

DR WPI: 2002-269010/31.

PT Identifying plant target proteins for herbicidally active compounds,
comprising aligning and comparing nucleic acid or amino acid sequences
from plant with nucleic acid or amino acid sequences from non-plant
organisms -

PS Claim 5; SEQ ID NO 2593; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
CC (ABR90790-ABR94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.

XX Sequence 354 AA;

Query Match 90.3%; Score 28; DB 23; Length 354;

Best Local Similarity 83.3%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 VRYEDL 6
Db 249 LRYEDL 254

RESULT 33
AAG44205
ID AAG44205 standard; Protein: 359 AA.

AC AAG44205;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 55343.

KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.

OS Arabidopsis thaliana.

PN EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127452.

PR 06-APR-1999; 99US-0128274.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139859.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140355.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 02-JUL-1999; 99US-0142154.
 PR 06-JUL-1999; 99US-0142055.
 PR 08-JUL-1999; 99US-0142390.
 PR 09-JUL-1999; 99US-0142820.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144864.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145148.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 03-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147935.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149472.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151303.
 PR 01-SEP-1999; 99US-0151338.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157553.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0158293.
 PR 13-OCT-1999; 99US-0158294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160815.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 359;
 Best Local Similarity 83.3%; Pred. No. 5.1e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 254 LYREDL 259

RESULT 34

AA646836
ID AA646836 standard; Protein: 359 AA.
XX
AC AA646836;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 58965.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EPI033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-012180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126784.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132853.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145085.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148584.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159631.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 22-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161820.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161933.
PR 29-OCT-1999; 99US-0162142.

Query Match 90.3%; Score 28; DB 21; Length 359;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 254 LRYEDL 259

RESULT 35
AAB49722
ID AAB49722 standard; Protein; 359 AA.
XX
AC AAB49722;
XX
DT 09-APR-2001 (first entry)
XX

DE Hydroxyjasmonic acid sulfotransferase AtST2a protein.
XX
KW Hydroxyjasmonic acid sulfotransferase; AtST2a; flowering time; cabbage;
KW jasmonate; genetically modified plant; lettuce; sugar cane; carrot;
KW increase vegetative growth; biomass increase.
XX
OS Arabidopsis thaliana.
XX
PN W0200102589-A2.
XX
PD 11-JAN-2001.
XX
PF 06-JUL-2000; 2000MO-CA00801.
XX
PR 06-JUL-1999; 99CA-2274873.
XX
PA (VARI/) VARIN L.
XX (GIBD/) GIDDA S.
PI Varin L, Gidda S;
PI
DR WPI: 2001-159272/16.
DR N-PSDB; AAF29177.
XX
XX Methods for modulating flowering in plants, particularly useful for
PT plants used in the food-processing industry, involves modifying the
PT endogenous level of compounds of the jasmonate family -
PS
PS Claim 50; Fig 8; 50pp: English.
XX
XX This invention relates to a method for modulating flowering in a plant.
CC The method comprises modifying the endogenous level of at least one
CC compound of the jasmonate family in a plant. The methods are used to
CC produce plants which are genetically modified to flower early or tardily
CC when compared to a corresponding plant that is not genetically modified,
CC where the modified plant has an increased (flower early) or lowered
CC (flower tardily) level of jasmonic acid, or a compound of the jasmonate
CC family. The method is useful for modulating flowering, particularly for
CC plants that are used in the food-processing industry and plants with
CC horticultural value. The method is particularly useful for e.g. delaying
CC flowering time in crops like lettuce, cabbage, sugar cane or carrots
CC which results in increased vegetative growth and biomass. The present
CC sequence represents the Arabidopsis thaliana AtST2a protein which is an
CC hydroxyjasmonic acid sulfotransferase. The AtST2a gene can be used in
CC the method of the invention.
CC
SQ Sequence 359 AA;
XX

Query Match 90.3%; Score 28; DB 22; Length 359;
Best Local Similarity 83.3%; Pred. No. 5.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 254 LRYEDL 259

RESULT 36
AA939919
ID AA939919 standard; Protein; 388 AA.
XX
AC AA939919;
XX
DT 08-DEC-1999 (first entry)
XX
DE Mouse glycosyl sulfotransferase-3 protein sequence.
XX
KW Glycosyl sulfotransferase; GST-3; detection; diagnosis; leukocyte homing;
KW selectin binding interaction; inflammation; lymphocyte homing; mouse;
KW secondary lymph organ.
XX
OS Mus sp.
XX

PN W09949018-A1.
 XX 30-SEP-1999.
 PD 26-FEB-1999; 99WO-US04316.
 XX 20-MAR-1998; 98US-0045284.
 PR 12-NOV-1998; 98US-0190911.
 XX (RESC) UNIV CALIFORNIA.
 PA (SYNT) SYNTAX USA INC.
 XX Bistrup A, Rosen SD, Tangemann K, Hemmerich S;
 PI WPI: 1999-580442/49.
 DR N-PSDB; AA220793.
 XX Human and murine glycosyl sulfotransferase 3 and related
 PT polynucleotides
 PS Claim 2; Fig 4; 59pp: English.
 XX This sequence is the mouse glycosyl sulfotransferase-3 (GST-3) of
 CC the invention. The nucleic acid sequences, probes and primers derived
 CC from these, proteins and antibodies are useful in detecting homologues.
 CC The sequences, antibodies and methods are useful in the diagnosis and
 CC treatment of diseases associated with selectin binding interactions,
 CC including conditions associated with or resulting from the homing of
 CC leukocytes to sites of inflammation and the normal homing of lymphocytes
 CC to secondary lymph organs.
 CC
 SQ Sequence 388 AA;
 Query Match 90.3%; Score 28; DB 20; Length 388;
 Best Local Similarity 83.3%; Pred. No. 5.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VREEDL 6
 :|||||
 DB 273 LREYDL 278
 RESULT 37
 AAM06480
 ID AAM06480 standard; Protein: 458 AA.
 XX
 AC AAM06480;
 XX
 DT 25-FEB-1997 (first entry)
 XX
 DE Chick chondroitin 6-sulphotransferase.
 XX
 KW Chondroitin 6-sulphotransferase; C6ST; chondroitin sulphate;
 KW proteoglycan.
 XX
 OS Gallus sp.
 XX
 XX Location/Qualifiers
 FH 1..33
 FT Key
 FT Peptide
 FT Protein
 FT Protein
 FT Protein
 FT Protein
 FT Protein
 FT Domain
 FT Modified-site
 FT /Label= Glycosylation

FT /note= "potential N-glycosylation site"
 FT 73..75
 FT /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 95..97
 FT /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 236..238
 FT /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 399..401
 FT /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT 443..445
 FT /label= Glycosylation
 FT /note= "potential N-glycosylation site"
 FT
 XX EP745668-A2.
 PN
 XX 04-DEC-1996.
 PD
 XX 30-MAY-1996; 96EP-0303687.
 PF
 XX 31-MAY-1995; 95JP-0134358.
 PR
 XX (SEK) SEIRAGAKU KOGYO CO LTD.
 PA
 XX Fukuta M, Habuchi O;
 PI WPI: 1997-013692/02.
 DR
 DR N-PSDB; AAT45037.
 XX
 PT DNA encoding chondroitin 6-sulpho:transferase - for recombinant
 PT production of C6ST, for use in industrial processes
 XX
 PS Claim 2; Page 18-20; 30pp: English.
 XX
 CC Chick embryo chondroitin 6-sulphotransferase (C6ST)
 CC (AAM06480) is capable of catalysing the transfer of a sulphate group
 CC from 3'-phosphadenosine 5' phosphosulphate to the hydroxyl group
 CC at the C-6 position of the N-acetylgalactosamine residue of
 CC chondroitin. Its amino acid sequence was deduced from a cDNA clone
 CC (AAT45037) isolated from a chick embryo chondrocyte cDNA library.
 CC Isolation of the cDNA allows the mass prodn. of C6ST in transformed
 CC host cells. Recombinant C6ST polypeptides are useful for analysing
 CC the activities of chondroitin sulphate and for modifying its
 CC function, and for raising antibodies.
 CC
 SQ Sequence 458 AA;
 Query Match 90.3%; Score 28; DB 18; Length 458;
 Best Local Similarity 83.3%; Pred. No. 6.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VREEDL 6
 :|||||
 DB 348 VREEDV 353
 RESULT 38
 AAB96675
 ID AAB96675 standard; Protein: 458 AA.
 XX
 AC AAB96675;
 XX
 DT 29-OCT-2001 (first entry)
 XX
 DE Putative P. abyssal Zn protease #5.
 XX
 KW Hyperthermophilic archaeon; hyperthermophilic protein.
 XX
 OS Pyrococcus abyssi.
 XX

PN FR2792651-A1.
 XX 27-OCT-2000.
 XX
 XX 21-APR-1999; 99FR-0005034.
 XX
 XX 21-APR-1999; 99FR-0005034.
 XX
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX (IFRE-) IFREMER INST FR RECH EXPL MER.
 XX
 XX Forterre P, Thierry JC, Prieur D, Dietrich J, Lecompte O;
 XX Querrelou J, Weissenbach J, Saurin W, Hellig R;
 XX WPI; 2001-126236/14.
 XX
 XX New nucleotide sequences isolated from Pyrococcus abyssi encode
 XX proteins useful in industry -
 XX
 XX Claim 7; Pages 1431-1433; 1657pp; French.
 XX
 XX The present invention relates to the genomic sequence of Pyrococcus
 XX abyssi (see AAF86431 and AAH41223-7) and P. abyssi proteins. P. abyssi is
 XX a hyperthermophilic archaeon, which is isolated from deep-sea
 XX hydrothermal vents. The present sequence is one such P. abyssi protein.
 XX The proteins of the present invention have various potential industrial
 XX uses, since the proteins are stable at very high temperatures, some up to
 XX 110 degrees centigrade.
 XX Note: This patent is in the same patent family as WO2000065062, which
 XX contains additional sequences as shown in AAB99132-AAB99143,
 XX AAH75903-AAH75920 and AAG66436.
 XX
 XX Sequence 458 AA;
 XX
 XX Query Match 90.3%; Score 28; DB 22; Length 458;
 XX Best Local Similarity 83.3%; Pred. No. 6.7e+02;
 XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 VRYEDL 6
 XX :|||||
 XX 25 LRYEDL 30
 XX
 XX RESULT 39
 XX AAW52863
 XX ID AAW52863 standard; Protein; 479 AA.
 XX
 XX AAW52863;
 XX
 XX 30-JUN-1998 (first entry)
 XX
 XX Glycosaminoglycan sulphotransferase protein.
 XX
 XX Glycosaminoglycan sulphotransferase; C6ST; N-acetyl-galactosamine;
 XX galactose; glycosaminoglycan; Chondroitin; keratan sulphate; diagnostic.
 XX
 XX Homo sapiens.
 XX
 XX EP821066-A1.
 XX
 XX 28-JAN-1998.
 XX
 XX 22-JUL-1997; 97EP-0305476.
 XX
 XX 24-JUL-1996; 96JP-0195063.
 XX
 XX (SEKG) SEIKAGAKU KOGYO CO LTD.
 XX
 XX Fukuta M, Habuchi O;
 XX WPI; 1998-102622/10.
 XX
 XX Human glycosaminoglycan sulphotransferase protein - catalyses

PT 6-sulphation of chondroitin to produce chondroitin sulphate for use
 XX in pharmaceutical(s)
 XX
 XX Claim 3; Page 21-23; 27pp; English.
 XX
 XX The human glycosaminoglycan sulphotransferase (C6ST) is a type II
 XX membrane protein derived from foetal brain tissue with a molecular weight
 XX of 50-55 kD. The protein transfers sulphate groups from a donor to the
 XX N-acetyl-galactosamine residue or galactose residue of glycosaminoglycan.
 XX C6ST displays substrate specificity transferring the sulphate group to
 XX the hydroxyl group position at C-6 of the N-acetyl-galactosamine residue
 XX of chondroitin and the hydroxyl group position at C-6 of the galactose
 XX residue of keratan sulphate. The protein is useful for studying the
 XX function of chondroitin sulphate and can provide chondroitin sulphate
 XX that may be useful in pharmaceuticals. Both the protein and DNA might
 XX be useful for treating or diagnosing diseases attributable to low C-6
 XX sulphation of chondroitin N-acetyl-galactosamine residues.
 XX
 XX Sequence 479 AA;
 XX
 XX Query Match 90.3%; Score 28; DB 19; Length 479;
 XX Best Local Similarity 83.3%; Pred. No. 7.1e+02;
 XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX 1 VRYEDL 6
 XX :|||||
 XX 369 VRYEDV 374
 XX
 XX RESULT 40
 XX AAR66031
 XX ID AAR66031 standard; Protein; 567 AA.
 XX
 XX AAR66031;
 XX
 XX 28-JUN-1995 (first entry)
 XX
 XX Mammalian MEK kinase (MEK 3).
 XX
 XX MEK kinase; MEK3; mitogen-activated protein kinase regulator;
 XX MAPK; cell atrophy inhibitor; Parkinson's; Alzheimer's; cancer;
 XX autoimmune diseases; allergies; wound healing; oncogenes;
 XX tumour agents; neurotrophic growth factor.
 XX
 XX Mus musculus.
 XX
 XX Key Location/Qualifiers
 XX FT 1..174
 XX FT /note="amino terminus regulatory domain"
 XX FT 175..179
 XX FT /note="regulatory hinge region"
 XX FT 357..567
 XX FT /note="catalytic domain"
 XX
 XX WO9424159-A.
 XX
 XX 27-OCT-1994.
 XX
 XX 15-APR-1994; 94WO-US04178.
 XX
 XX 15-APR-1993; 93US-0049254.
 XX
 XX (NAJDE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 XX
 XX Johnson GL;
 XX WPI; 1994-357747/44.
 XX
 XX New MEK kinase protein and related antibodies and nucleic acid
 XX regulator of mitogen activated protein kinase, useful
 XX therapeutically to inhibit cell atrophy, to screen for oncogenes
 XX etc.

PS Claim 6; Page 13; 84pp; English.

XX AA079325 encodes AAR66029 the mammalian MEK kinase (MEK 1), other
 CC unique mammalian MEK kinases identified by PCR are described in
 CC AAR66030 (MEK 2), AAR66031 (MEK 3) and AAR66032 (MEK 4). MEK is an
 CC activator, independent of Raf protein, of mitogen-activated protein
 CC kinases (MAPK). Inactivation of MEK can be used in the treatment
 CC of some cancers, autoimmune diseases and allergies, while
 CC stimulation can promote wound healing. MEK can also be used to
 CC alleviate cellular atrophy in Parkinson's or Alzheimer's by acting
 CC as a neurotrophic growth factor, and to screen for oncogenes and
 CC tumour agents.

SQ Sequence 567 AA;

Query Match 90.3%; Score 28; DB 15; Length 567;
 Best Local Similarity 83.3%; Pred. No. 8.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 DB 64 VRYEDV 69

RESULT 41
 ABB92573
 ID ABB92573 standard; Protein; 612 AA.

XX AC ABB92573;
 XX 31-MAY-2002 (first entry)
 DE Herbicidally active polypeptide SEQ ID NO 1784.
 XX Herbicidal; plant; agriculture; herbicide.
 KW Arabidopsis thaliana.
 OS WO200210210-A2.
 PN 07-FEB-2002.
 PD 28-AUG-2001; 2001WO-EP09892.
 XX 28-AUG-2001; 2001WO-EP09892.
 PR 28-AUG-2001; 2001WO-EP09892.
 XX (FARB) BAYER AG.
 PA Tietjen K, Weidner M;
 XX Tietjen K, Weidner M;
 PT WPI; 2002-269010/31.
 DR Identifying plant target proteins for herbicidally active compounds,
 XX comprising aligning and comparing nucleic acid or amino acid sequences
 PT from plant with nucleic acid or amino acid sequences from non-plant
 PT organisms -
 PS Claim 5; SEQ ID NO 1784; 261pp + Sequence Listing; English.

XX The invention relates to identifying target proteins
 CC (ABB90790-ABB94016) for herbicidally active compounds, comprising
 CC aligning and comparing nucleic acid or amino acid sequences from plant
 CC with nucleic acid or amino acid sequences from non-plant organisms using
 CC suitable search parameters, where plant sequences having an E-value
 CC greater by a factor of 3 than the E-value of most similar non-plant
 CC sequences are selected. The polypeptides or nucleic acids encoding them
 CC are useful for identifying modulators. The identified modulators are
 CC useful as herbicides.

SQ Sequence 612 AA;

Query Match 90.3%; Score 28; DB 23; Length 612;
 Best Local Similarity 83.3%; Pred. No. 9.3e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 :|||||
 DB 230 LRYEDL 235

RESULT 42
 AAR77546
 ID AAR77546 standard; Protein; 626 AA.

XX AC AAR77546;
 XX 12-MAR-1996 (first entry)
 DE MEK3 protein.
 XX MEK3; mitogen ERK kinase kinase; signal transduction; homeostasis;
 KW cancer; inflammation; neurological disorder; autoimmune disease;
 KW allergy; hormone-related disease.
 XX Mus sp.
 OS Mus sp.
 XX Key
 XX Location/Qualifiers
 FT Domain 1..174
 FT /label= Regulatory_domain
 FT Region 175..179
 FT /label= Regulatory_hinge_sequence
 FT Domain 357..626
 FT /label= Catalytic_domain

XX WO9528421-A1
 XX 26-OCT-1995.
 PD 14-OCT-1994; 94WO-US11690.
 PF 15-APR-1994; 94WO-US04178.
 XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
 PA Johnson GL;
 XX Johnson GL;
 PI WPI; 1995-373762/48.
 XX N-PSDB; AAT05572.
 DR New signal-regulated kinase proteins and nucleic acids - used for
 XX regulating cell responsiveness for treating e.g. tumours,
 PT auto-immune disease, inflammation or neuronal disorders.
 PS Claim 1; Page 31-34; 147pp; English.

XX MEK2, MEK3, MEK4 and MEK5 (AAR77545-48) are mitogen ERK kinase
 CC kinases capable of regulating the activity of mitogen-activated protein
 CC kinase. The proteins can be produced in heterologous host cells by
 CC expression of cDNAs (AAT05571-74) isolated from PC12 and HL60 cells.
 CC They are capable of regulating signal transduction in cells. By
 CC modulating the activity of an MEK-dependent pathway relative to that of
 CC the Raf-dependent pathway it is possible to inhibit cell growth for
 CC treatment of cancer, autoimmunity, etc.

SQ Sequence 626 AA;

Query Match 90.3%; Score 28; DB 16; Length 626;
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 |||||
 DB 64 VRYEDV 69

RESULT 43

AAW56159
ID AAW56159 standard; Protein; 626 AA.
XX
AC AAW56159;
XX
DT 17-JUL-1998 (first entry)
XX
DE A mitogen-activated protein kinase kinase (MAPKK).
XX
KW Mitogen-activated protein kinase kinase; MAPKK; mouse;
KW extracellular signal-regulated kinase kinase; MEKK; regulation;
KW signal transduction; raf-independent arm; screening assay; treatment;
KW disorder; cancer; autoimmune disease; inflammation; allergy;
KW neuronal disease; Parkinson's disease; Alzheimer's disease; ds.
XX
OS Mammalia.
XX
PN US5753446-A.
XX
PD 19-MAY-1998.
XX
PF 06-JUN-1995; 95US-0472934.
XX
PR 15-APR-1993; 93US-0049254.
PR 14-OCT-1994; 94US-0323460.
PR 21-FEB-1995; 95US-0354516.
PR 12-MAY-1995; 95US-0440421.
XX
PA (NAE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Johnson GL;
XX
PT WPI; 1998-311395/27.
DR N-PSDB; AAW22678.
XX
XX Screening assay for regulators of MEKK signal transduction - using
PT mammalian MEKK polypeptide
XX
PS Claim 13; Columns 51-54; 48pp; English.
XX
CC The present sequence represents a mitogen-activated protein kinase
CC kinase (MAPKK) (also known as extracellular signal-regulated kinase
CC kinase (MEKK)). The protein, which is serine/threonine kinase is capable
CC of regulating signal transduction in cells. It regulates the activity of
CC elements of the raf-independent arm of MEKK. A screening assay for
CC compounds that regulate signal transduction by a MEKK protein comprises
CC contacting a reaction mixture containing a mammalian MEKK polypeptide and
CC a test compound and determining the effect of the test compound on an
CC indicator of signal transduction by the MEKK polypeptide in the reaction
CC mixture. Compounds identified by the above assay can be used to prepare
CC therapeutic compositions for treating disorders that are subject to
CC regulation or cure by manipulating a signal transduction pathway in
CC cells involved in the disorders, e.g. cancer, autoimmune diseases,
CC inflammations, allergies, and neuronal diseases such as Parkinson's
CC disease and Alzheimer's disease.
XX
SQ Sequence 626 AA;

Query Match 90.3%; Score 28; DB 19; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
|||||
DB 64 VRYEDV 69

RESULT 44
ID AAW43320
XX AAW43320 standard; Protein; 626 AA.
AC AAW43320;
XX

DT 24-JAN-2000 (first entry)
XX
XX Mitogen ERK kinase kinase, MEKK-3, protein sequence.
DE
XX
KW Mitogen ERK kinase kinase; MEKK; MEKK-3; neurological disorder; cancer;
KW extracellular signal-regulated kinase; inflammation; autoimmune disease;
KW allergic reaction; hormone related disease; therapy.
XX
XX Homo sapiens.
OS
XX
PN US5981265-A.
XX
PD 09-NOV-1999.
XX
PF 05-JUN-1995; 95US-0461146.
XX
PR 15-APR-1993; 93US-0049254.
PR 12-MAY-1995; 95US-0440421.
PR 15-APR-1994; 94MO-US04178.
PR 14-OCT-1994; 94US-0323460.
PR 14-OCT-1994; 94MO-US11690.
PR 28-NOV-1994; 94US-0345516.
XX
PA (NAE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
PI Johnson GL;
XX
PT WPI; 1999-633328/54.
DR N-PSDB; AAW31879.
XX
XX Regulating mitogen extracellular signal-regulated kinase kinase protein
PT activity, useful for the treatment of cancer, neurological diseases and
XX autoimmune diseases -
XX
PS Claim 2; Column 65-70; 94pp; English.
XX
CC This sequence is the mitogen ERK (extracellular signal-regulated
CC kinase) kinase kinase-3 (MEKK-3). The invention relates to a method of
CC regulating MEKK protein activity in a fungal cell by transforming or
CC transfecting the cell with a nucleic acid encoding an MEKK protein. The
CC MEKK protein is useful for treating cancer, inflammation, neurological
CC disorders, autoimmune diseases, allergic reactions, and hormone related
CC diseases.
XX
SQ Sequence 626 AA;

Query Match 90.3%; Score 28; DB 20; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
|||||
DB 64 VRYEDV 69

RESULT 45
ID AAW42106
XX AAW42106 standard; Protein; 626 AA.
AC AAW42106;
XX
DT 09-DEC-1999 (first entry)
XX
DE Human MEKK3 protein sequence.
XX
KW MEKK1; MEKK2; MEKK3; mitogen-activated protein kinase; MAPK; ERK;
KW extracellular regulated kinase; signal transduction; regulation;
KW MAPK/ERK; MEK; MEKK; inflammation; cellular proliferation;
KW differentiation; development; cell death.
XX
OS Homo sapiens.
XX
PN WO9947686-A2.

```

XX 23-SEP-1999.
XX
XX 15-MAR-1999; 99WO-US05556.
XX
XX 16-MAR-1998; 98US-0078153.
XX
XX 04-SEP-1998; 98US-0099165.
XX
XX (CADU-) CADUS PHARM CORP.
XX
XX Johnson GL;
XX
XX WPI; 1999-571843/48.
XX
XX N-PSDB; AA225071.
XX
XX New human MEK polynucleotides and polypeptides, used for regulating
XX signal transduction in cells -
XX
XX Claim 12; Fig 10; 159pp; English.
XX
XX The present sequence represents human mitogen-activated protein kinase/
XX extracellular response kinase (MAPK/ERK) kinase kinase (MEKK),
XX specifically designated MEKK3. The MEKK proteins are used to modulate
XX and regulate signal transduction in cells, as well as for regulation of
XX gene transcription in a cell encoding MEKK, where the cell is involved
XX in inflammation, regulation of cellular proliferation and
XX differentiation, regulation of development, regulation of cell death or
XX regulation of inflammation. They are also used to prepare antibodies.
XX MEKK polynucleotides can be used to produce the protein recombinantly
XX and as a source of probes and primers.
XX
XX Sequence 626 AA:
XX
XX Query Match 90.3%; Score 28; DB 20; Length 626;
XX Best Local Similarity 83.3%; Pred. No. 9.5e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VRYEDL 6
XX
XX 64 VRYEDV 69
XX
XX RESULT 46
XX AAY42109
XX ID AAY42109 standard; Protein: 626 AA.
XX
XX AC AAY42109;
XX
XX 09-DEC-1999 (first entry)
XX
XX Murine MEKK3 protein sequence.
XX
XX MEKK1; MEKK2; MEKK3; mitogen-activated protein kinase; MAPK; ERK;
XX extracellular regulated kinase; signal transduction; regulation;
XX MAPK/ERK; MEK; MKK; Inflammation; cellular proliferation;
XX differentiation; development; cell death.
XX
XX Mus musculus.
XX
XX WO9947686-A2.
XX
XX 23-SEP-1999.
XX
XX 15-MAR-1999; 99WO-US05556.
XX
XX 16-MAR-1998; 98US-0078153.
XX
XX 04-SEP-1998; 98US-0099165.
XX
XX (CADU-) CADUS PHARM CORP.
XX
XX Johnson GL;
XX
XX WPI; 1999-571843/48.
XX

```

```

DR N-PSDB; AA225074.
XX
XX New human MEK polynucleotides and polypeptides, used for regulating
XX signal transduction in cells -
XX
XX Example 3; Fig 12; 159pp; English.
XX
XX The present sequence represents murine mitogen-activated protein kinase/
XX extracellular response kinase (MAPK/ERK) kinase kinase (MEKK),
XX specifically designated MEKK3. The MEKK proteins are used to modulate
XX and regulate signal transduction in cells, as well as for regulation of
XX gene transcription in a cell encoding MEKK, where the cell is involved
XX in inflammation, regulation of cellular proliferation and
XX differentiation, regulation of development, regulation of cell death or
XX regulation of inflammation. They are also used to prepare antibodies.
XX MEKK polynucleotides can be used to produce the protein recombinantly
XX and as a source of probes and primers.
XX
XX Sequence 626 AA:
XX
XX Query Match 90.3%; Score 28; DB 20; Length 626;
XX Best Local Similarity 83.3%; Pred. No. 9.5e+02;
XX Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 VRYEDL 6
XX
XX 64 VRYEDV 69
XX
XX RESULT 47
XX AAW73533
XX ID AAW73533 standard; Protein: 626 AA.
XX
XX AC AAW73533;
XX
XX 04-MAR-1999 (first entry)
XX
XX MEKK3 protein.
XX
XX Mitogen activated protein kinase kinase; MEKK; MAPK; MEK; cancer;
XX apoptosis regulator; autoimmune disease; inflammation; allergy; therapy;
XX neuronal disorder.
XX
XX Mammalia.
XX
XX US5854043-A.
XX
XX 29-DEC-1998.
XX
XX 14-OCT-1994; 94US-0323460.
XX
XX 14-OCT-1994; 94US-0323460.
XX
XX 15-APR-1993; 93US-0049254.
XX
XX 15-APR-1994; 94WO-US04178.
XX
XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.
XX
XX Johnson GL;
XX
XX WPI; 1999-094912/08.
XX
XX N-PSDB; AAY45616.
XX
XX Mitogen activated protein kinase kinase kinases and their fragments -
XX used for regulating signalling from growth factor receptors, e.g. to
XX modulate apoptosis for treatment of cancer, autoimmune disease and
XX inflammation
XX
XX Claim 5; Column 21-26; 96pp; English.
XX
XX This sequence is the MEKK3 protein of the invention. MEKK proteins
XX are mitogen-activated protein kinase (MAPK) kinase (MEK) kinase proteins.
XX MEKKs phosphorylate and activate MEK proteins and other signal
XX transduction molecules, so can regulate signalling initiated from a

```


CC growth factor receptor in a way different from that involving Raf
CC protein. Particularly MEKs, or their fragments, are involved in
CC regulation of apoptosis so they, or agents that increase their activity,
CC are used to treat cancers, autoimmune diseases, inflammation, allergies,
CC neuronal disorders (e.g. Alzheimer's or Parkinson's diseases) and in
CC wound healing. MEKs are also useful for identifying agents that
CC regulate signal transduction from cell surface receptors (e.g. from their
CC effect on ability of MEK to phosphorylate a substrate such as MEK or Jun
CC extracellular signal-regulated kinase).

SO Sequence 626 AA:

Query Match 90.3%; Score 28; DB 20; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||:
DB 64 VREEDV 69

RESULT 48

ID AAB01218 standard; Protein; 626 AA.

AC AAB01218;

DT 19-DEC-2000 (first entry)

DE Murine MEK3.

KW Mouse; MEK3; mitogen ERK kinase kinase; signal transduction pathway;

KW apoptosis; cancer; autoimmune disease; inflammatory response;

KW allergic response; neuronal disorder; Parkinson's disease;

KW Alzheimer's disease.

OS Mus sp.

PN US6074861-A.

PD 13-JUN-2000.

XX 05-JUN-1995; 95US-0461145.

PR 15-APR-1993; 93US-0049254.

PR 12-MAY-1995; 93US-0440421.

PR 15-APR-1994; 94WO-0504178.

PR 14-OCT-1994; 94US-0323460.

PR 14-OCT-1994; 94WO-0511690.

PR 21-FEB-1995; 95US-0354516.

XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX Johnson GL;

PI MPI; 2000-411281/35.

DR N-PSDB; AAA49224.

XX Novel mitogen extracellular signal-regulated kinase kinase (MEK3)

PT protein useful for treating cancer, inflammation, autoimmune diseases,

PT neurological disorders and hormone related disease in animals

PS Claim 1; Column 69-74; 92pp; English.

XX The present sequence is the murine mitogen ERK kinase kinase (MEK3)

CC protein sequence. MEK3 is involved in a signal transduction pathway

CC which can ultimately lead to apoptosis. The proteins regulated by MEK3

CC include the MEK and MAPK proteins and c-Myc. Its coding sequence was

CC identified by amplifying genomic DNA with a similar sequence to

CC the MEK1 gene. The gene and protein can be used to treat diseases such

CC as cancer, autoimmune disease, inflammatory responses, allergic responses

CC and neuronal disorders including Parkinson's disease and Alzheimer's

CC disease.

XX Sequence 626 AA;

Query Match 90.3%; Score 28; DB 21; Length 626;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||:
DB 64 VREEDV 69

RESULT 49
ID AAM48938 standard; Protein; 626 AA.

AC AAM48938;

DT 19-APR-2002 (first entry)

DE Murine MEK3.

KW Mouse; MEK3; mitogen ERK kinase kinase; enzyme; cancer; neuroprotective;

KW autoimmune disease; signal transduction; allergy; inflammation;

KW neurological disorder; hormone-related disease; apoptosis; infection;

KW cytosolic; immunosuppressive; antiinflammatory; antiallergic;

KW nootropic; antiparkinsonian; contraceptive.

OS Mus musculus.

PN US6333170-B1.

PD 25-DEC-2001.

XX 05-APR-1996; 96US-0628829.

PR 15-APR-1993; 93US-0049254.

PR 14-OCT-1994; 94US-0323460.

PR 12-MAY-1995; 95US-0440421.

PR 06-JUN-1995; 95US-0472934.

XX (NAJE-) NAT JEWISH CENT IMMUNOLOGY & RESPIRATORY.

XX Johnson GL;

PI MPI; 2002-163179/21.

DR N-PSDB; AAL41581.

XX New isolated nucleic acid encoding mitogen extracellular

PT signal-regulated kinase kinase, useful for gene therapy of e.g. cancer

PT and for recombinant protein production

PS Claim 10; Column 151-154; 125pp; English.

XX The present invention provides the protein and coding sequences of a

CC number of murine mitogen extracellular signal-regulated kinase (ERK)

CC kinase kinase (MEK) enzymes. The sequences can be used to treat a wide

CC range of diseases including cancer, autoimmune diseases, inflammation,

CC allergies, degenerative neurological diseases and hormone-related

CC diseases, and for inhibiting spermatogenesis or oocyte maturation for

CC contraception. The present sequence is the murine MEK3 protein

CC sequence.

SO Sequence 626 AA;

Query Match 90.3%; Score 28; DB 23; Length 626;

Best Local Similarity 83.3%; Pred. No. 9.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||:
DB 64 VREEDV 69

RESULT 50

ID ABB68414 standard; Protein; 793 AA.

AC ABB68414;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 32034.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US09231.

PR 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656660/75.

DR N-PSDB; ABL12517.

PS Disclosure: SEQ ID NO 32034; 21pp + Sequence listing; English.

CC The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB101840-AB16175) and the encoded proteins sequences (AB101840-AB16175).

CC The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SQ Sequence 793 AA;

Query Match 90.3%; Score 28; DB 22; Length 793;

Best Local Similarity 83.3%; Pred. No. 1.2e+03;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VAREDL 6

DB 673 LRYEDL 678

Search completed: February 20, 2003, 13:32:28

Job time : 35.6286 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:11 ; Search time 9.6 Seconds
(without alignments)
60.084 Million cell updates/sec

Title: US-09-816-825-9
Perfect score: 31
Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	31	100.0	281	2 B90167	conserved hypothet
2	31	100.0	428	2 G75453	hypothetical prote
3	31	100.0	484	2 JE0261	N-acetylglucosamin
4	31	100.0	631	2 A64578	conserved hypothet
5	31	100.0	748	2 T30634	hypothetical prote
6	31	100.0	973	2 T21463	hypothetical prote
7	31	100.0	6260	2 T30228	polyketide synthas
8	30	96.8	562	2 JC7628	glucoside 3-dehydr
9	30	96.8	579	2 H87451	oxidoreductase, GM
10	30	96.8	1322	2 B71440	hypothetical prote
11	29	93.5	234	2 H87385	hypothetical prote
12	29	93.5	355	2 F96020	probable iron ABC
13	29	93.5	393	2 T08203	polygalacturonase
14	29	93.5	405	2 D84871	probable polygalac
15	29	93.5	430	2 G64446	ATP-dependent 26S
16	29	93.5	1401	2 T39225	MAP kinase kinase
17	29	93.5	1416	2 D71350	probable DNA-dirac
18	28	90.3	195	1 B69441	conserved hypothet
19	28	90.3	233	2 C86849	regulatory protein
20	28	90.3	273	2 T42928	immediate-early pr
21	28	90.3	277	2 C69608	chitosanase csn -
22	28	90.3	306	2 A83634	hypothetical prote
23	28	90.3	333	2 A84523	probable steroid s
24	28	90.3	337	2 C38453	anaerobic sulfite
25	28	90.3	337	2 A60825	anaerobic sulfite
26	28	90.3	346	2 T29003	hypothetical prote
27	28	90.3	388	2 G70729	hypothetical prote
28	28	90.3	436	2 AH2447	molybdopterin bios
29	28	90.3	458	2 C75180	tlld related prote

30	28	90.3	458	2 A57397	Chondroitin 6-sulf
31	28	90.3	484	2 JC7350	N-acetylglucosamin
32	28	90.3	486	2 JC7351	N-acetylglucosamin
33	28	90.3	503	2 F90169	hypothetical prote
34	28	90.3	507	2 E89804	alkyl hydroperoxid
35	28	90.3	612	2 T45641	beta-D-glucan exch
36	28	90.3	1107	2 B91271	probable periplasm
37	28	90.3	1107	2 B86112	probable periplasm
38	28	90.3	1107	2 E65226	hypothetical 123.8
39	28	90.3	2233	1 ZLN2P3	genome polyprotein
40	27	87.1	73	2 H83960	hypothetical prote
41	27	87.1	108	2 F72649	hypothetical prote
42	27	87.1	133	2 H69433	response regulator
43	27	87.1	160	2 A25571	myosin ICI catalyt
44	27	87.1	200	2 C95315	hypothetical prote
45	27	87.1	203	2 C95194	Multi/nudix family
46	27	87.1	203	2 H98060	mutator protein [1
47	27	87.1	225	2 G70340	Mg(2+) transport A
48	27	87.1	225	2 S28679	glycoprotein 30 -
49	27	87.1	265	2 C90026	hypothetical prote
50	27	87.1	273	2 S07014	hypothetical prote
51	27	87.1	273	2 JN0037	hypothetical 32k P
52	27	87.1	283	2 E81895	probable prolipopr
53	27	87.1	307	2 E95934	probable enzyme, C
54	27	87.1	309	2 JC5697	placental transfor
55	27	87.1	311	2 D69881	yabc protein homol
56	27	87.1	312	2 B72708	probable signal re
57	27	87.1	332	2 A70772	probable sulfate A
58	27	87.1	334	2 C64130	oxidoreductase BH3
59	27	87.1	341	2 G82014	dtDPglucose 4,6-de
60	27	87.1	346	2 S47045	dtDPglucose 4,6-de
61	27	87.1	352	2 H83975	hypothetical prote
62	27	87.1	355	2 G81242	dtDPglucose 4,6-de
63	27	87.1	356	2 S49011	yolk protein 2 - b
64	27	87.1	360	2 S42431	dtDPglucose 4,6-de
65	27	87.1	386	2 F88694	protein skn-1 (imp
66	27	87.1	396	2 F89850	3-phosphoglycerate k
67	27	87.1	428	2 A69085	3-isopropylmalate
68	27	87.1	433	2 B69435	aconitase (acn) ho
69	27	87.1	439	2 H83842	diaminopimelate de
70	27	87.1	439	2 C55072	Guanine deaminase
71	27	87.1	439	2 D91098	hypothetical prote
72	27	87.1	439	2 H85943	hypothetical prote
73	27	87.1	465	2 A31102	glutamyl-tRNA amid
74	27	87.1	465	2 E98184	probable amidase P
75	27	87.1	474	2 B82227	exodeoxyribonuclea
76	27	87.1	476	2 H82177	conserved hypothet
77	27	87.1	488	2 D86698	SOS response unuc
78	27	87.1	488	2 T43124	conserved hypothet
79	27	87.1	499	2 S43324	zeta-carotene desa
80	27	87.1	499	2 AG2509	zeta-carotene desa
81	27	87.1	504	2 H87422	methy-accepting c
82	27	87.1	514	2 B64055	peptidase-like oligop
83	27	87.1	520	2 T21462	hypothetical prote
84	27	87.1	525	2 A83013	conserved hypothet
85	27	87.1	527	2 F69378	conserved hypothet
86	27	87.1	533	2 A42143	skn-1 - Caenorhabd
87	27	87.1	554	2 T51213	hypothetical prote
88	27	87.1	557	2 B86466	hypothetical prote
89	27	87.1	561	2 E98192	probable oxidoredu
90	27	87.1	561	2 A83094	oxidoreductase Atu
91	27	87.1	562	2 A86773	hypothetical prote
92	27	87.1	605	2 G85651	probable membrane
93	27	87.1	605	2 D90791	probable membrane
94	27	87.1	617	2 G70039	conserved hypothet
95	27	87.1	621	2 B71457	probable phosphoen
96	27	87.1	628	2 A84930	glucose inhibited
97	27	87.1	630	2 H97562	hypothetical prote
98	27	87.1	630	2 AH2783	peptidyl-prolyl ci
99	27	87.1	631	2 E71933	hypothetical prote
100	27	87.1	653	2 T03102	semaphorin homolog
101	27	87.1	666	2 AH1167	ABC transporter, A
102	27	87.1	685	2 F96032	hypothetical prote

103	27	87.1	694	2	D86615	176	26	83.9	686	2	D71292	probable DNA recom
104	27	87.1	694	2	G72009	177	26	83.9	686	2	B96526	unknown protein (i
105	27	87.1	716	2	T37710	178	26	83.9	693	1	B64139	DNA helicase recog
106	27	87.1	747	2	T42599	179	26	83.9	721	2	F87611	tonB-dependent rec
107	27	87.1	753	1	WZBER8	180	26	83.9	727	1	S17854	NADH2 dehydrogenas
108	27	87.1	759	2	E87443	181	26	83.9	728	2	S45403	hypothetical prote
109	27	87.1	769	1	WZBE54	182	26	83.9	760	2	G71417	hypothetical prote
110	27	87.1	769	2	B87681	183	26	83.9	762	2	UC7174	N,N-dimethylformam
111	27	87.1	773	2	T46283	184	26	83.9	799	2	F83456	xanthine dehydroge
112	27	87.1	785	2	B72608	185	26	83.9	830	2	F83288	conserved hypotet
113	27	87.1	809	2	H87455	186	26	83.9	849	2	E86306	Similar to tutei
114	27	87.1	891	2	T40417	187	26	83.9	886	1	UC5085	replication licens
115	27	87.1	926	2	T04679	188	26	83.9	902	1	S42228	replication licens
116	27	87.1	976	2	A87319	189	26	83.9	904	1	T10067	hypothetical prote
117	27	87.1	1060	2	F88710	190	26	83.9	975	2	AC2517	protein F5011.4 (i
118	27	87.1	1079	2	T30996	191	26	83.9	975	2	A86258	CYRTR
119	27	87.1	1094	2	S49313	192	26	83.9	1057	1	OYMSAR	atrial natriuretic
120	27	87.1	1096	1	S61917	193	26	83.9	1057	1	TS7963	natriuretic peptid
121	27	87.1	1137	2	T19414	194	26	83.9	1057	2	TS7963	guanylyl cyclase A
122	27	87.1	1174	2	T43081	195	26	83.9	1057	2	TS5319	hypothetical prote
123	27	87.1	1352	2	G71051	196	26	83.9	1166	2	T15628	hypothetical prote
124	27	87.1	1359	2	T10235	197	26	83.9	1226	2	AB3327	hypothetical mem
125	27	87.1	1393	2	E72122	198	26	83.9	1396	2	G71529	DNA-directed RNA p
126	27	87.1	1393	2	B86501	199	26	83.9	1396	2	F81686	DNA-directed RNA p
127	27	87.1	1397	2	B81548	200	26	83.9	1397	2	T46354	hypothetical prote
128	27	87.1	1704	2	T43141	201	26	83.9	1450	2	A84780	probable ABC trans
129	27	87.1	3433	1	S28381	202	26	83.9	2548	2	E59435	myosin IXA (import
130	27	87.1	10223	2	T30225	203	26	83.9	2626	2	T31099	myosin-Rhocap prot
131	26	83.9	105	2	F90503	204	26	83.9	3144	2	S64791	Vps13 protein - ye
132	26	83.9	121	2	S35234	205	26	80.6	72	2	T49026	ubiquinol-cytochro
133	26	83.9	136	2	C81436	206	25	80.6	88	2	T07369	ubiquinol-cytochro
134	26	83.9	166	2	T49694	207	25	80.6	97	2	S57628	ribosomal protein
135	26	83.9	166	2	T64740	208	25	80.6	102	2	A71126	thymidylate syntha
136	26	83.9	201	2	F97063	209	25	80.6	114	2	AH3529	hypothetical prote
137	26	83.9	231	2	T23530	210	25	80.6	117	1	W0BP37	flav protein (limp
138	26	83.9	246	2	C70316	211	25	80.6	122	2	T05357	gene 0.3 protein -
139	26	83.9	253	2	D95121	212	25	80.6	133	2	B82977	conserved hypotet
140	26	83.9	253	2	H97990	213	25	80.6	137	2	T41575	ubiquinol-cytochro
141	26	83.9	261	2	S10321	214	25	80.6	145	2	F83267	conserved hypotet
142	26	83.9	261	2	S75926	215	25	80.6	147	2	F77310	hypothetical prote
143	26	83.9	284	2	T22671	216	25	80.6	162	2	D69899	conserved hypotet
144	26	83.9	288	1	B48583	217	25	80.6	163	2	AC1939	peptide methionine
145	26	83.9	302	2	D83479	218	25	80.6	163	2	H96014	conserved hypotet
146	26	83.9	306	2	T29990	219	25	80.6	169	2	A54776	probable membrane
147	26	83.9	339	2	T28178	220	25	80.6	169	2	G90692	probable membrane
148	26	83.9	344	2	A95053	221	25	80.6	185	2	C85543	conserved hypotet
149	26	83.9	356	2	F97923	222	25	80.6	200	2	H97513	nucol (A7245398) (
150	26	83.9	356	2	AC2907	223	25	80.6	200	2	AC2732	NADH ubiquinone ox
151	26	83.9	365	2	B69220	224	25	80.6	203	2	AC3611	formylmethionine d
152	26	83.9	378	2	G97682	225	25	80.6	203	2	AC3290	transcription regu
153	26	83.9	384	2	G82976	226	25	80.6	235	2	AD1735	hypothetical prote
154	26	83.9	408	2	T40698	227	25	80.6	238	2	A69463	2-hydroxy-6-oxo-6-
155	26	83.9	413	2	T08297	228	25	80.6	241	2	B97164	sigma factor of St
156	26	83.9	430	2	S73475	229	25	80.6	248	2	S38956	epithelial cell ma
157	26	83.9	448	2	S45112	230	25	80.6	258	2	H72269	ABC transporter, A
158	26	83.9	448	1	E71033	231	25	80.6	264	2	G75045	probable zinc-bind
159	26	83.9	465	2	A31810	232	25	80.6	286	2	A75583	hypothetical prote
160	26	83.9	488	2	B96521	233	25	80.6	286	2	T25719	hypothetical prote
161	26	83.9	498	2	S12061	234	25	80.6	300	1	S16815	SNP1 protein - yea
162	26	83.9	507	2	C82901	235	25	80.6	305	2	G95122	ABC transporter, A
163	26	83.9	526	1	P5XR15	236	25	80.6	306	2	E83200	probable transcrip
164	26	83.9	526	1	P5XR15	237	25	80.6	310	2	A33489	hypothetical prote
165	26	83.9	526	1	A45185	238	25	80.6	316	1	YK5AT3	hypothetical prote
166	26	83.9	526	1	A45185	239	25	80.6	318	2	C89920	thymidylate syntha
167	26	83.9	527	1	S18762	240	25	80.6	321	2	H71729	hypothetical prote
168	26	83.9	527	1	S18762	241	25	80.6	326	2	H82987	maeonesium/cobalt t
169	26	83.9	596	2	G97133	242	25	80.6	331	2	D65515	hypothetical prote
170	26	83.9	602	2	S36475	243	25	80.6	331	2	T25184	
171	26	83.9	615	2	H82635	244	25	80.6				
172	26	83.9	633	2	F70283	245	25	80.6				
173	26	83.9	662	2	E86509	246	25	80.6				
174	26	83.9	662	2	B72114	247	25	80.6				
175	26	83.9	671	2	T23271	248	25	80.6				
					E82355							

249	25	80.6	334	2	S74737	conserved hypotnet	322	25	80.6	664	2	T12968	hypothetical prote
250	25	80.6	346	2	B69350	conserved hypotnet	323	25	80.6	716	2	T37830	probable glutamate
251	25	80.6	346	2	E70904	hypothetical prote	324	25	80.6	727	2	A33552	NADH2 dehydrogenas
252	25	80.6	347	2	AC2840	alcohol dehydrogen	325	25	80.6	736	2	T05137	protein kinase hom
253	25	80.6	350	2	C64359	endoglucanase homo	326	25	80.6	778	2	JC7797	scallop unconventi
254	25	80.6	350	2	D70805	hypothetical prote	327	25	80.6	787	2	C75058	probable beta-gala
255	25	80.6	351	2	F90409	GIP binding conser	328	25	80.6	801	2	AE1950	hypothetical prote
256	25	80.6	357	2	A97618	alcohol dehydrogen	329	25	80.6	866	2	C97662	hypothetical prote
257	25	80.6	363	2	S75088	probable GTP-bind	330	25	80.6	866	2	AE2886	conserved hypotnet
258	25	80.6	365	2	S74847	hypothetical prote	331	25	80.6	873	2	T25442	hypothetical prote
259	25	80.6	366	2	S66016	probable GTP-bind	332	25	80.6	883	1	RNBp17	DNA-directed RNA p
260	25	80.6	370	2	C69309	conserved hypotnet	333	25	80.6	884	1	RNBp13	DNA-directed RNA p
261	25	80.6	372	2	F70467	hypothetical prote	334	25	80.6	897	2	T43628	phosphatidylinosit
262	25	80.6	374	2	AE2227	transposase alr337	335	25	80.6	899	2	S12319	pre-mRNA splicing
263	25	80.6	375	1	A23689	limulus clotting e	336	25	80.6	928	2	S50578	hypothetical prote
264	25	80.6	377	2	S21302	succinate dehydrog	337	25	80.6	929	2	T35683	ftsK homolog - Str
265	25	80.6	385	2	S46532	polygalacturonase	338	25	80.6	971	2	H71719	hypothetical prote
266	25	80.6	394	2	F97178	spore coat polysac	339	25	80.6	976	2	H97269	zn-dependent pepti
267	25	80.6	400	2	T25605	hypothetical prote	340	25	80.6	979	2	J00894	Pil5 protein - Myc
268	25	80.6	402	2	T45518	hypothetical prote	341	25	80.6	979	2	B80601	hypothetical prote
269	25	80.6	402	2	AC2304	transposase alr388	342	25	80.6	988	2	S77211	iso-leucine-tRNA li
270	25	80.6	402	2	AC2352	transposase alr801	343	25	80.6	1013	2	B75383	hypothetical prote
271	25	80.6	402	2	AD2506	transposase alr722	344	25	80.6	1052	2	T00067	hypothetical prote
272	25	80.6	402	2	AC2506	transposase alr723	345	25	80.6	1095	2	PC1114	SKDC25 protein -
273	25	80.6	402	2	AI2318	transposase alr410	346	25	80.6	1133	2	T01920	probable RNA-direc
274	25	80.6	402	2	H86649	rhannosyltransfera	347	25	80.6	1249	2	S54376	tripeptidyl-peptid
275	25	80.6	403	2	AE2518	transposase alr732	348	25	80.6	1249	2	S68431	tripeptidyl-peptid
276	25	80.6	403	2	AI2363	transposase alr446	349	25	80.6	1262	2	I48855	tripeptidyl-peptid
277	25	80.6	403	2	AB1951	transposase alr115	350	25	80.6	1286	2	T18734	hypothetical prote
278	25	80.6	403	2	B95175	TPR domain protein	351	25	80.6	1302	2	T23236	hypothetical prote
279	25	80.6	409	1	EEEGT	translation elonga	352	25	80.6	1450	2	T45888	ABC transporter-11
280	25	80.6	413	2	AG2456	transposase alr520	353	25	80.6	1465	2	S31262	TYB protein - years
281	25	80.6	414	2	C98041	conserved hypotnet	354	25	80.6	1467	2	PC1253	TYB protein - years
282	25	80.6	425	2	H90415	hypothetical prote	355	25	80.6	1490	2	UC5145	DNA (cytosine-5')-
283	25	80.6	436	2	D84782	probable proline t	356	25	80.6	1537	2	UC4172	DNA (cytosine-5')-
284	25	80.6	439	2	T47713	proline transporte	357	25	80.6	1676	2	E71410	probable centromer
285	25	80.6	439	2	T50688	proline transport	358	25	80.6	1680	2	T01367	hypothetical prote
286	25	80.6	441	2	T50688	proline transport	359	25	80.6	1690	2	S41467	DNA-directed RNA p
287	25	80.6	442	2	T50692	proline transport	360	25	80.6	1748	2	S63127	probable membrane
288	25	80.6	442	2	T50687	proline transport	361	25	80.6	1802	2	S52611	TYB protein - years
289	25	80.6	442	2	C82823	type I restriction	362	25	80.6	1803	2	S68894	ribosomal protein
290	25	80.6	444	2	T20600	hypothetical prote	363	25	80.6	1819	2	H82331	conserved hypotnet
291	25	80.6	450	2	G83536	hypothetical prote	364	25	80.6	1819	2	T12008	hypothetical prote
292	25	80.6	452	2	A70389	L-seeryl-tRNAse se	365	25	80.6	17428	2	T17428	FS506 polypeptide s
293	25	80.6	456	2	T19817	hypothetical prote	366	25	80.6	73	2	E72588	hypothetical prote
294	25	80.6	471	2	AC3622	mannose-1-phosphat	367	25	80.6	83	2	E64482	hypothetical 8.2K
295	25	80.6	473	2	T50690	proline transport	368	25	80.6	86	2	T12077	late embryogenesis
296	25	80.6	481	2	T01850	UMP-glucose glucos	369	25	80.6	89	2	D82298	ribosomal protein
297	25	80.6	488	2	F71825	replicative DNA he	370	25	80.6	91	2	H82331	conserved hypotnet
298	25	80.6	488	2	T09734	1-aminocyclopropan	371	25	80.6	100	2	S69860	hypothetical prote
299	25	80.6	488	2	B64690	replicative DNA he	372	25	80.6	102	2	JC2343	heat shock protein
300	25	80.6	491	2	E69368	glutamine syntheta	373	25	80.6	102	2	JC2344	heat shock protein
301	25	80.6	498	2	B75373	GGDEF family prote	374	25	80.6	102	2	A72739	probable DNA-direc
302	25	80.6	500	2	T16630	hypothetical prote	375	25	80.6	110	1	VHBPDL	major capsid prote
303	25	80.6	520	1	S50990	galactokinase (EC	376	25	80.6	110	2	B80833	major capsid prote
304	25	80.6	532	2	S54571	probable membrane	377	25	80.6	110	2	G90900	major capsid prote
305	25	80.6	535	2	A84464	hypothetical prote	378	25	80.6	114	2	G89966	probable head deco
306	25	80.6	536	2	F90299	acylaminoacyl-pept	379	25	80.6	115	2	D85690	protein T08A9.6 [1
307	25	80.6	548	2	S27958	transcription fact	380	25	80.6	115	2	B70950	hypothetical prote
308	25	80.6	550	2	H71952	ATP-dependent zinc	381	25	80.6	117	2	G90464	hypothetical prote
309	25	80.6	550	2	F64555	cell division prot	382	25	80.6	121	2	D48234	hypothetical prote
310	25	80.6	558	2	D88163	protein F10C1.2a [383	25	80.6	125	2	T38537	probable single-st
311	25	80.6	573	2	S45903	hypothetical prote	384	25	80.6	125	2	S46670	bulious pemphigoid
312	25	80.6	575	2	AD2361	flavoprotein (limp	385	25	80.6	126	2	C97616	hypothetical prote
313	25	80.6	585	2	S48929	hypothetical prote	386	25	80.6	126	2	A12838	hypothetical prote
314	25	80.6	588	2	H89885	succinate dehydrog	387	25	80.6	127	2	AB2252	aspartate 1-decarb
315	25	80.6	589	2	S46326	hypothetical prote	388	25	80.6	127	2	B89862	conserved hypotnet
316	25	80.6	595	2	H87350	dehydratase, ilvd/	389	25	80.6	129	2	A30798	lysine carboxypept
317	25	80.6	603	2	T04733	auxin-regulated pr	390	25	80.6	130	2	H72722	hypothetical prote
318	25	80.6	605	2	S36592	El protein - human	391	25	80.6	131	2	AE1586	transcription regu
319	25	80.6	609	2	S36481	El protein - human	392	25	80.6	132	2	B84295	hypothetical prote
320	25	80.6	616	2	S64746	ORC protein - yea	393	25	80.6	139	2	T33713	conserved hypotnet
321	25	80.6	620	2	S64304	hypothetical prote	394	25	80.6	141	2	T28189	hypothetical prote

395	24	77.4	142	2	B86552	glycerol-3-phospha
396	24	77.4	147	2	T35563	ribosomol protein
397	24	77.4	150	2	T36372	probable asnc-fam1
398	24	77.4	151	2	C90048	conserved hypothet
399	24	77.4	153	2	T16012	hypothetical prote
400	24	77.4	156	2	A41509	myosin essential I
401	24	77.4	156	2	UT0902	chaperonin 60 beta
402	24	77.4	157	2	A25183	myosin essential I
403	24	77.4	158	2	H87643	conserved hypothet
404	24	77.4	164	2	A81335	probable signal-tr
405	24	77.4	165	2	G69304	transcription regu
406	24	77.4	165	2	G83415	hypothetical prote
407	24	77.4	169	2	T21917	3'-aminoacylcoside
408	24	77.4	170	2	T40612	NADH2 dehydrogenas
409	24	77.4	172	2	JEO381	hypothetical prote
410	24	77.4	174	2	B75176	chaperonin 60 beta
411	24	77.4	177	2	PS0374	hypothetical prote
412	24	77.4	177	2	S71014	probable atpd prot
413	24	77.4	177	2	C70679	hypothetical prote
414	24	77.4	183	2	A84537	hypothetical prote
415	24	77.4	184	2	T28333	ORF MSV172 hypote
416	24	77.4	185	2	E70445	C-terminal fumarat
417	24	77.4	188	2	E72860	ORF-84 protein - A
418	24	77.4	190	1	KLS8A	calcium-binding pr
419	24	77.4	190	2	A83102	conserved hypothet
420	24	77.4	190	2	S48062	NADH2 dehydrogenas
421	24	77.4	190	2	T03231	NADH2 dehydrogenas
422	24	77.4	192	1	T03231	calcium-binding pr
423	24	77.4	192	2	S62130	NADH2 dehydrogenas
424	24	77.4	192	2	T11930	NADH2 dehydrogenas
425	24	77.4	193	2	D64769	YajB protein - Esc
426	24	77.4	193	2	G90685	probable glycoprot
427	24	77.4	193	2	C85356	probable glycoprot
428	24	77.4	193	2	H90474	hypothetical prote
429	24	77.4	194	2	T48071	hypothetical prote
430	24	77.4	196	2	A41163	Ca2+/calmodulin-de
431	24	77.4	197	2	H87491	NADH dehydrogenase
432	24	77.4	199	2	A32183	tropomyosin TPM1 -
433	24	77.4	199	2	S75334	hypothetical prote
434	24	77.4	200	2	H98209	hypothetical prote
435	24	77.4	200	2	A13076	conserved hypothet
436	24	77.4	200	2	F90098	putative ribosom
437	24	77.4	201	2	F70468	conserved hypothet
438	24	77.4	203	2	G69898	conserved hypothet
439	24	77.4	204	2	S78165	NADH2 dehydrogenas
440	24	77.4	208	2	F90223	conserved hypothet
441	24	77.4	210	2	G97233	hypothetical prote
442	24	77.4	212	2	AC1055	peptide methionine
443	24	77.4	212	2	S25996	hypothetical prote
444	24	77.4	213	2	B84333	isopentenyl pyroph
445	24	77.4	213	2	C70943	hypothetical prote
446	24	77.4	215	1	ADEGFP	L-fuculose-phospha
447	24	77.4	215	2	D91086	L-fuculose-1-phosp
448	24	77.4	215	2	F85931	N-acylhomoserine l
449	24	77.4	216	2	A10120	probable heme expo
450	24	77.4	218	2	AF0333	hypothetical prote
451	24	77.4	227	2	T49725	hypothetical prote
452	24	77.4	228	2	T16678	hypothetical prote
453	24	77.4	232	2	AD0071	probable DedA-fam1
454	24	77.4	232	2	AD0372	hypothetical prote
455	24	77.4	233	2	AB1366	protein gp33 [Bact
456	24	77.4	233	2	A13590	DNA directed DNA p
457	24	77.4	234	2	A40710	CD30 ligand - huma
458	24	77.4	235	2	B72606	gp53 protein - Myc
459	24	77.4	237	2	AE2037	hypothetical prote
460	24	77.4	238	2	C70479	thiol-disulfide in
461	24	77.4	241	2	B83907	hypothetical prote
462	24	77.4	242	2	A71972	hypothetical prote
463	24	77.4	242	2	C64537	conserved hypothet
464	24	77.4	243	2	T13105	minor tail protein
465	24	77.4	243	2	S67186	hypothetical prote
466	24	77.4	244	2	A14114	merr-family transc
467	24	77.4	245	2	S00413	photosystem II oxy
468	24	77.4	245	2	S68691	neutrophil granule
469	24	77.4	245	2	S03888	photosystem II oxy
470	24	77.4	248	2	T02122	hypothetical prote
471	24	77.4	248	2	C83865	hypothetical prote
472	24	77.4	248	2	G97340	acyl-ACP thioester
473	24	77.4	249	2	E84072	hypothetical prote
474	24	77.4	251	2	D69363	conserved hypothet
475	24	77.4	252	2	T03873	photosystem II oxy
476	24	77.4	253	2	H90656	probable membrane
477	24	77.4	253	2	H85507	probable membrane
478	24	77.4	254	2	T02873	probable photostst
479	24	77.4	254	2	F90408	conserved hypothet
480	24	77.4	255	2	AC0437	probable membrane
481	24	77.4	255	2	D86693	conserved hypothet
482	24	77.4	257	1	G65206	NADH pyrophosphata
483	24	77.4	257	2	G91243	hypothetical prote
484	24	77.4	257	2	E86091	hypothetical prote
485	24	77.4	257	2	AD0932	heat-labile entero
486	24	77.4	258	1	OLECA	photosystem II oxy
487	24	77.4	258	1	F210X2	photosystem II oxy
488	24	77.4	258	2	S22763	photosystem II oxy
489	24	77.4	259	2	J50771	photosystem II oxy
490	24	77.4	259	2	A75547	hypothetical prote
491	24	77.4	259	2	A72116	biotin apo-protein
492	24	77.4	259	2	A86507	biotin protein lig
493	24	77.4	260	2	S10016	photosystem II oxy
494	24	77.4	260	2	T32440	photosystem II oxy
495	24	77.4	262	2	J70390	hypothetical prote
496	24	77.4	265	1	YAZGF7	type II site-speci
497	24	77.4	265	2	S63532	S-antigen precurs
498	24	77.4	267	2	S00005	NAD(P)H-quinone ox
499	24	77.4	267	2	S17446	photosystem II oxy
500	24	77.4	268	2	T00088	photosystem II oxy
501	24	77.4	269	2	F86482	ABC-type transport
502	24	77.4	271	1	PRECT9	protein F5J5.7 (Im
503	24	77.4	271	2	JU0103	kanamycin kinase
504	24	77.4	271	2	JU0145	kanamycin kinase
505	24	77.4	271	2	S16630	APH protein - Salm
506	24	77.4	274	2	H83525	kanamycin kinase
507	24	77.4	277	2	E86833	conserved hypothet
508	24	77.4	278	2	T03942	transcription regu
509	24	77.4	279	2	S75709	rRNA N-glycosidase
510	24	77.4	281	2	S24723	hypothetical prote
511	24	77.4	282	2	B82323	tryptophan synthas
512	24	77.4	287	2	S39072	dimethyladenosine
513	24	77.4	291	2	C97742	NADH2 dehydrogenas
514	24	77.4	291	2	B90256	hypothetical prote
515	24	77.4	291	2	S74914	conserved hypothet
516	24	77.4	291	2	G71300	probable protein-m
517	24	77.4	292	2	T03085	ribonuclease homol
518	24	77.4	292	2	T03122	hypothetical prote
519	24	77.4	293	2	S51571	hypothetical prote
520	24	77.4	293	2	E75366	glutanyl-tRNA synt
521	24	77.4	295	2	A53372	lyser-type regulato
522	24	77.4	296	2	E69025	conserved hypothet
523	24	77.4	297	2	C70251	hypothetical prote
524	24	77.4	297	2	C70231	hypothetical prote
525	24	77.4	300	2	T36255	hypothetical prote
526	24	77.4	301	1	RLZMR1	rRNA N-glycosidase
527	24	77.4	301	2	F82287	transcription regu
528	24	77.4	302	2	S11860	rRNA N-glycosidase
529	24	77.4	303	2	G83093	UDP-3-O-acyl-N-ace
530	24	77.4	304	2	S11859	rRNA N-glycosidase
531	24	77.4	306	2	A83239	conserved hypothet
532	24	77.4	309	1	YAKON7	S-antigen precurs
533	24	77.4	309	2	F71976	probable [acyl-car
534	24	77.4	309	2	B64531	probable [acyl-car
535	24	77.4	312	2	F90540	hypothetical prote
536	24	77.4	316	2	T36133	cysteine synthase
537	24	77.4	319	2	G70373	Mg(2+) and Co(2+)
538	24	77.4	319	2	D90482	dehydrogenase, pro
539	24	77.4	320	2	H87443	NADP-dependent qui
540	24	77.4	321	2	A42507	F5L protein - vacc

541	24	77.4	322	2	E36213	614	24	77.4	411	2	I52966	myosin heavy chain
542	24	77.4	327	2	T34720	615	24	77.4	415	2	T32995	hypothetical prote
543	24	77.4	328	2	AE0885	616	24	77.4	414	2	B75070	cytochrome-c3 hydr
544	24	77.4	333	2	T37145	617	24	77.4	416	2	T12781	IMPB/MUCB/SAMB fam
545	24	77.4	334	2	AD321	618	24	77.4	417	2	E64200	serine-tRNA ligase
546	24	77.4	334	2	T18983	619	24	77.4	418	2	A11784	UV-damage repair p
547	24	77.4	336	2	A72322	620	24	77.4	420	2	C71407	hypothetical prote
548	24	77.4	339	2	F71060	621	24	77.4	422	2	C82666	conserved hypotet
549	24	77.4	340	2	T28347	622	24	77.4	424	2	S15146	hypothetical prote
550	24	77.4	342	2	S50455	623	24	77.4	426	2	E84164	neutral proteinase
551	24	77.4	343	2	H90867	624	24	77.4	428	2	S77064	glutamate dehydrog
552	24	77.4	343	2	A85751	625	24	77.4	429	2	T32290	hypothetical prote
553	24	77.4	343	2	AB0662	626	24	77.4	432	2	T11652	probable translio
554	24	77.4	343	2	E64882	627	24	77.4	432	2	C84351	histidyl-tRNA synt
555	24	77.4	343	2	C97656	628	24	77.4	433	1	S48515	adenylosuccinate s
556	24	77.4	345	2	B84342	629	24	77.4	433	2	T17654	hypothetical prote
557	24	77.4	346	2	T31760	630	24	77.4	438	2	T45602	glucosyltransferas
558	24	77.4	347	2	E89788	631	24	77.4	443	2	S03257	methyl coenzyme M
559	24	77.4	348	2	AB1496	632	24	77.4	443	2	A87793	cytochrome p450 fa
560	24	77.4	348	2	T21254	633	24	77.4	444	2	E87368	UDP-N-acetylmuramo
561	24	77.4	348	2	T28467	634	24	77.4	445	2	DB1777	hypothetical prote
562	24	77.4	348	2	I36839	635	24	77.4	445	2	S67147	probable diphospha
563	24	77.4	350	2	H96768	636	24	77.4	445	2	F70190	glucosyltransferas
564	24	77.4	351	2	A84827	637	24	77.4	448	1	T45605	T64 protein precu
565	24	77.4	352	2	JC7781	638	24	77.4	449	2	S07714	replicative DNA he
566	24	77.4	354	2	JC4212	639	24	77.4	449	2	T45605	clusterin - qual
567	24	77.4	354	2	T17561	640	24	77.4	450	2	AG1081	hypothetical prote
568	24	77.4	355	2	I40819	641	24	77.4	450	2	AG1081	conserved hypotet
569	24	77.4	355	2	D97278	642	24	77.4	451	2	AI1438	lysine carboxypept
570	24	77.4	355	2	T00102	643	24	77.4	452	2	I50131	probable aldehyde
571	24	77.4	355	2	A12879	644	24	77.4	455	2	E83260	hexokinase (EC 2.7
572	24	77.4	356	2	A89807	645	24	77.4	458	2	BA6733	hexokinase (EC 2.7
573	24	77.4	358	2	D89823	646	24	77.4	458	2	S02074	glucokinase (EC 2.7
574	24	77.4	359	2	H64547	647	24	77.4	462	2	T36807	penicillin amidase
575	24	77.4	359	2	E71960	648	24	77.4	464	2	A6157	hexokinase (EC 2.7
576	24	77.4	360	1	S30109	649	24	77.4	465	2	A46157	hexokinase (EC 2.7
577	24	77.4	360	2	T22576	650	24	77.4	465	2	I49693	glucokinase (EC 2.7
578	24	77.4	361	2	H83884	651	24	77.4	465	2	A83536	penicillin amidase
579	24	77.4	365	2	T46849	652	24	77.4	466	2	BA6157	hexokinase (EC 2.7
580	24	77.4	365	2	T20652	653	24	77.4	466	2	A55163	nucleolar protein
581	24	77.4	366	2	G70230	654	24	77.4	468	2	A57176	glutamate synthase
582	24	77.4	367	2	T18839	655	24	77.4	470	2	T25578	hypothetical prote
583	24	77.4	370	2	E70354	656	24	77.4	471	2	B84412	hypothetical prote
584	24	77.4	374	2	D24827	657	24	77.4	472	1	WMADP6	early E1B 52K prot
585	24	77.4	375	2	C24827	658	24	77.4	472	2	S17595	cannabinoic recept
586	24	77.4	375	2	B24827	659	24	77.4	473	2	A33117	glutamate synthase
587	24	77.4	375	2	T24132	660	24	77.4	473	2	B67885	hypothetical prote
588	24	77.4	375	2	JS0427	661	24	77.4	473	2	E67885	hypothetical prote
589	24	77.4	377	2	AC2336	662	24	77.4	485	2	C64680	hypothetical prote
590	24	77.4	379	2	C55205	663	24	77.4	491	2	B95662	hypothetical prote
591	24	77.4	381	2	B83072	664	24	77.4	492	2	T18221	chromosome condens
592	24	77.4	385	2	T12980	665	24	77.4	495	2	G87684	histidyl-tRNA synt
593	24	77.4	389	2	T30068	666	24	77.4	497	2	P83651	hypothetical prote
594	24	77.4	391	2	S18666	667	24	77.4	497	2	S57073	probable membrane
595	24	77.4	391	2	AD1653	668	24	77.4	499	2	E84149	transposase (25) B
596	24	77.4	396	2	C68283	669	24	77.4	500	1	E69291	4-hydroxyphenylace
597	24	77.4	397	1	SS1594	670	24	77.4	505	2	S57672	hypothetical prote
598	24	77.4	398	2	H95057	671	24	77.4	506	2	C81704	monooxygenase-rela
599	24	77.4	398	2	A97927	672	24	77.4	508	2	G90830	hypothetical prote
600	24	77.4	398	2	E64741	673	24	77.4	508	2	DB5668	unknown protein en
601	24	77.4	398	2	G90650	674	24	77.4	508	2	E64786	Yok protein - Bsc
602	24	77.4	398	2	G85501	675	24	77.4	509	2	A40448	DNA-binding protei
603	24	77.4	398	2	AE0529	676	24	77.4	510	2	A90204	conserved hypotet
604	24	77.4	399	1	A55578	677	24	77.4	511	2	T50910	pyruvate dehydroge
605	24	77.4	400	2	S23888	678	24	77.4	515	2	S52321	penton base protei
606	24	77.4	402	1	E70302	679	24	77.4	516	2	T33061	hypothetical prote
607	24	77.4	403	2	E82810	680	24	77.4	517	2	A48250	UAF-homologous pr
608	24	77.4	405	1	BA2606	681	24	77.4	518	2	B83472	hypothetical prote
609	24	77.4	405	2	S46828	682	24	77.4	520	2	DB9870	peptide chain rele
610	24	77.4	406	2	S68866	683	24	77.4	521	2	S06218	colicin E1 - shige
611	24	77.4	409	2	D64336	684	24	77.4	524	2	C81367	phosphoenolpyruvat
612	24	77.4	409	2	T28355	685	24	77.4	525	1	XZAD10	penton protein (II
613	24	77.4	410	2	T20397	686	24	77.4	529	2	T35831	probable D-3-phosp

687	24	77.4	529	2	T33738	hypothetical prote	760	24	77.4	727	2	G84487	probable DNA repli
688	24	77.4	530	2	T42520	conserved hypochet	761	24	77.4	728	1	HHCH90	heat shock protein
689	24	77.4	531	2	B55788	glutathioneS-transfer	762	24	77.4	730	2	D72210	conserved hypochet
690	24	77.4	536	2	S66716	glutamate-tRNA lig	763	24	77.4	732	1	S05238	peptidyl-dipeptida
691	24	77.4	542	2	H64339	hypothetical prote	764	24	77.4	732	1	HHHR86	heat shock protein
692	24	77.4	543	2	T33352	probable proteinase	765	24	77.4	732	1	H68395	hypothetical prote
693	24	77.4	543	2	S64850	probable membrane	766	24	77.4	733	1	HHMS86	hypothetical prote
694	24	77.4	544	1	S41389	penton protein (II	767	24	77.4	737	2	T46243	hypothetical prote
695	24	77.4	544	2	H87473	medium-chain-fatty	768	24	77.4	764	2	T45793	hypothetical prote
696	24	77.4	545	2	S27995	probable acid-CoA	769	24	77.4	766	2	T20003	hypothetical prote
697	24	77.4	549	2	JC2564	heat shock protein	770	24	77.4	775	1	A39631	replication licens
698	24	77.4	550	1	A34576	crystal protein pr	771	24	77.4	776	2	T48317	cyclin F - mouse
699	24	77.4	559	2	H89931	DNA repair protein	772	24	77.4	779	2	G87573	xanthine dehydrog
700	24	77.4	559	2	H82532	conserved hypochet	773	24	77.4	782	2	G96698	hypothetical prote
701	24	77.4	561	2	G96752	unknown protein P2	774	24	77.4	783	2	T37457	tricorn proteinase
702	24	77.4	565	2	H69342	GTP-binding protei	775	24	77.4	785	2	I50180	cadherin-7 - chick
703	24	77.4	569	2	A11347	hypothetical prote	776	24	77.4	786	2	A55501	hypothetical prote
704	24	77.4	572	2	T41371	hypothetical prote	777	24	77.4	788	2	T25967	hypothetical prote
705	24	77.4	581	2	T33737	hypothetical prote	778	24	77.4	789	1	A45617	replication licens
706	24	77.4	588	2	FW0007	hypothetical prote	779	24	77.4	790	2	I51638	cyclin F - human
707	24	77.4	589	2	H90112	chaperonin 62.5K b	780	24	77.4	790	2	E85726	6-phosphofructokin
708	24	77.4	589	2	S06954	intermediate filam	781	24	77.4	790	2	B64903	F-cadherin - Afri
709	24	77.4	590	2	D84523	probable auxin-reg	782	24	77.4	792	2	D90891	hypothetical prote
710	24	77.4	593	2	A48459	S antigen, heat sh	783	24	77.4	792	2	B82756	hypothetical prote
711	24	77.4	595	2	T06412	probable chaperon	784	24	77.4	805	2	B75515	organic solvent to
712	24	77.4	597	2	S74612	translation elonga	785	24	77.4	807	2	A64575	sensory box/GGDP
713	24	77.4	600	2	UT0901	chaperonin 60 beta	786	24	77.4	807	2	T02738	VilB4 homolog - He
714	24	77.4	600	2	B96597	Rubisco subunit b1	787	24	77.4	812	2	A53016	probable AAA-type
715	24	77.4	601	2	T02633	hypothetical prote	788	24	77.4	830	2	C86191	myosin heavy chain
716	24	77.4	604	2	H89914	hypothetical prote	789	24	77.4	838	2	F97768	hypothetical prote
717	24	77.4	607	2	H81676	lipid-A-disacchari	790	24	77.4	843	2	H86629	hypothetical prote
718	24	77.4	610	2	S41315	hypothetical prote	791	24	77.4	862	1	A49346	protein P22G5.10 l
719	24	77.4	613	2	T43933	DNA-directed DNA p	792	24	77.4	882	2	AC3082	aldehyde dehydrog
720	24	77.4	618	2	F87390	TonB-dependent rec	793	24	77.4	893	2	T03864	ferric aerobactin
721	24	77.4	618	2	A75015	phosphoenolpyruvat	794	24	77.4	901	2	D70116	hypothetical prote
722	24	77.4	619	2	I64087	translation elonga	795	24	77.4	908	2	C70168	transcription elon
723	24	77.4	620	2	H85431	ATPase-like protei	796	24	77.4	913	2	F82958	DNA polymerase I p
724	24	77.4	627	2	T21605	hypothetical prote	797	24	77.4	935	2	H84491	hypothetical prote
725	24	77.4	632	2	H70339	NADH2 dehydrogenas	798	24	77.4	936	2	T23393	hypothetical prote
726	24	77.4	637	2	H86693	DNA primase (EC 2.	799	24	77.4	946	1	IYHR2	inter-alpha-trypsi
727	24	77.4	638	2	D69957	conserved hypochet	800	24	77.4	960	2	A48083	chromosome segrega
728	24	77.4	638	2	AD1166	probable peptidogl	801	24	77.4	982	2	T18576	serine-threonine k
729	24	77.4	640	2	F90587	lipoprotein (Impor	802	24	77.4	1057	4	B47521	RNA-dependent RNA
730	24	77.4	642	2	JC2485	DNA primase (EC 2.	803	24	77.4	1077	2	S45395	hypothetical prote
731	24	77.4	644	2	T21137	hypothetical prote	804	24	77.4	1082	2	S37837	suppressor protein
732	24	77.4	648	2	G72279	ATP-dependent DNA	805	24	77.4	1108	2	T41188	probable ubiquitin
733	24	77.4	650	2	B87456	TonB-dependent rec	806	24	77.4	1108	2	AF1047	probable membrane
734	24	77.4	664	2	A72215	exonuclease ABC c	807	24	77.4	1112	2	D75036	cell division cont
735	24	77.4	664	2	I39467	bullous pemphigoid	808	24	77.4	1113	2	S48495	probable membrane
736	24	77.4	664	2	T17317	hypothetical prote	809	24	77.4	1184	2	D86387	probable protein p
737	24	77.4	668	1	F69794	DNA ligase (NAD) (810	24	77.4	1195	1	S26722	DNA-directed RNA p
738	24	77.4	678	2	H88187	protein C18H9.8 Y1	811	24	77.4	1195	2	E96615	hypothetical prote
739	24	77.4	679	2	T40219	hypothetical prote	812	24	77.4	1206	2	S24407	formin isoform IV
740	24	77.4	680	2	H81987	probable DNA helic	813	24	77.4	1213	2	A41724	limb deformity (Id
741	24	77.4	680	2	E81043	ATP-dependent DNA	814	24	77.4	1221	2	T30529	ubiquitin (ubiquit
742	24	77.4	686	1	D70172	DNA recombinase (r	815	24	77.4	1237	2	AC1583	ubiquitin carboxyl
743	24	77.4	688	2	S61249	probable vitron pr	816	24	77.4	1265	2	T06916	internalin protein
744	24	77.4	688	2	T48176	receptor like prot	817	24	77.4	1289	2	I84505	DNA-directed RNA p
745	24	77.4	692	2	H82041	ATP-dependent DNA	818	24	77.4	1306	1	A31759	calcium-dependent
746	24	77.4	693	1	JH0265	DNA recombinase (E	819	24	77.4	1316	1	H1BPD7	peptidyl-dipeptida
747	24	77.4	693	2	G91194	DNA helicase RecG	820	24	77.4	1331	1	XORPDH	internal vitron pr
748	24	77.4	693	2	AE0005	ATP-dependent DNA	821	24	77.4	1335	1	XOMSDB	xanthine dehydrog
749	24	77.4	693	2	AE0969	ATP-dependent DNA	822	24	77.4	1364	2	T10236	xanthine dehydrog
750	24	77.4	703	2	A84563	probable AAA-type	823	24	77.4	1390	2	A45455	nucleoporin 155 -
751	24	77.4	704	2	H86041	hypothetical prote	824	24	77.4	1413	2	G84790	probable ABC trans
752	24	77.4	712	2	B87683	nonB-dependent rec	825	24	77.4	1435	2	D96693	protein putative A
753	24	77.4	717	1	HHHR83	heat shock protein	826	24	77.4	1437	2	C75198	activator 1, repli
754	24	77.4	718	2	C82817	ATP-dependent DNA	827	24	77.4	1451	2	A36468	SPT protein - yea
755	24	77.4	719	2	E88504	protein B0361.8 [1	828	24	77.4	1468	2	S11515	formin - mouse
756	24	77.4	721	2	T40945	hypothetical prote	829	24	77.4	1469	2	H96622	probable ABC trans
757	24	77.4	724	1	HHH084	heat shock protein	830	24	77.4	1588	2	T38660	probable transcrip
758	24	77.4	724	1	HHMS84	heat shock protein	831	24	77.4	1688	2	D75489	hypothetical prote
759	24	77.4	725	2	JC1468	heat shock protein	832	24	77.4	1679	2	S48385	hypothetical prote

833	24	77.4	1736	2	A47747	tight junction pro	906	23	74.2	136	2	H71668	ribosomal protein
834	24	77.4	1792	2	T13939	myosin V - fruit f	907	23	74.2	136	2	B83935	hypothetical prote
835	24	77.4	1828	2	B59254	myosin heavy chain	908	23	74.2	136	2	S61014	hypothetical prote
836	24	77.4	1830	1	S19188	myosin-V - chicken	909	23	74.2	137	1	FXME	flavodoxin - Megs
837	24	77.4	1853	1	A46761	myosin heavy chain	910	23	74.2	137	2	AC1817	hypothetical prote
838	24	77.4	1855	2	A59254	transcription init	911	23	74.2	138	2	S36115	interferon - Japan
839	24	77.4	1865	1	I48155	gag-pol-like fusio	912	23	74.2	139	2	T01592	hypothetical prote
840	24	77.4	1870	2	C47521	hypothetical prote	913	23	74.2	140	2	AC2348	mannose-6-phosphat
841	24	77.4	1882	2	T00069	transcription init	914	23	74.2	141	2	AH2994	hypothetical prote
842	24	77.4	1893	1	A40262	Utrascriptlon init	915	23	74.2	141	2	B98289	hypothetical prote
843	24	77.4	2206	1	GNNY21	genome polypotein	916	23	74.2	141	2	B82859	hypothetical prote
844	24	77.4	2291	2	S11238	polymerase - Berne	917	23	74.2	143	2	T13303	hypothetical prote
845	24	77.4	2531	2	T16743	hypothetical prote	918	23	74.2	144	2	D89934	hypothetical prote
846	24	77.4	2649	2	A40937	bullous pemphigoid	919	23	74.2	145	2	A97483	hypothetical prote
847	24	77.4	2748	2	S57976	nuclear migration	920	23	74.2	146	2	D66740	hypothetical prote
848	24	77.4	2810	2	T23298	hypothetical prote	921	23	74.2	146	2	D62117	conserved hypotet
849	24	77.4	2871	2	A55567	fibillin I - boyl	922	23	74.2	151	2	B70862	hypothetical prote
850	24	77.4	3002	2	A47221	fibillin I - boyl	923	23	74.2	151	2	T21354	hypothetical prote
851	24	77.4	3268	2	S69625	hypothetical prote	924	23	74.2	153	2	H86862	hypothetical prote
852	24	77.4	3332	1	GNNVJS	genome polypotein	925	23	74.2	155	2	S71257	hypothetical prote
853	24	77.4	3432	1	GNNVJS	genome polypotein	926	23	74.2	155	2	C97885	major latex protei
854	24	77.4	3587	2	I40485	surfactin syntheta	927	23	74.2	155	2	D75022	hypothetical prote
855	24	77.4	4830	2	E68679	polyketide synthet	928	23	74.2	157	2	S75487	6-pyruvoyl tetrahy
856	24	77.4	7962	2	I38346	elastic titlin - hu	929	23	74.2	159	2	C84299	hypothetical prote
857	24	77.4	31	2	S49191	hypothetical prote	930	23	74.2	161	2	C72201	hypothetical prote
858	23	74.2	45	2	S78733	protein YMR158c-b	931	23	74.2	161	2	S72986	hypothetical prote
859	23	74.2	48	2	F86043	unknown protein en	932	23	74.2	161	2	T13688	deatlilase 2 homo
860	23	74.2	55	2	D33052	Flis protein - Rhi	933	23	74.2	162	2	B87655	conserved hypotet
861	23	74.2	55	2	B95344	hypothetical prote	934	23	74.2	164	2	JX0135	inorganic diphosph
862	23	74.2	58	2	S58628	hypothetical prote	935	23	74.2	164	2	G81077	hypothetical prote
863	23	74.2	58	2	T12949	hypothetical prote	936	23	74.2	164	2	D81193	hypothetical prote
864	23	74.2	61	2	H91196	hypothetical prote	937	23	74.2	164	2	G90661	probable peptide c
865	23	74.2	61	2	G69257	hypothetical prote	938	23	74.2	166	2	B42509	E/R protein - vaco
866	23	74.2	63	1	O0BC1M	yaia protein - Esc	939	23	74.2	166	2	E83512	probable peptide c
867	23	74.2	63	2	AD0550	conserved hypotet	940	23	74.2	166	2	E64748	translation releas
868	23	74.2	63	2	C85534	hypothetical prote	941	23	74.2	166	2	C71068	hypothetical prote
869	23	74.2	63	2	G90683	hypothetical prote	942	23	74.2	167	2	AG0100	general secretion
870	23	74.2	64	2	T36824	hypothetical prote	943	23	74.2	167	2	S67993	hypothetical prote
871	23	74.2	71	2	JQ0277	hypothetical BK pr	944	23	74.2	168	2	S65242	hypothetical prote
872	23	74.2	72	2	AH3218	hypothetical cytos	945	23	74.2	168	2	T36980	hypothetical prote
873	23	74.2	85	2	S31018	gene 73 protein -	946	23	74.2	171	2	S65487	light-harvesting c
874	23	74.2	87	1	R3NW15	ribosomal protein	947	23	74.2	172	2	A72237	hypothetical prote
875	23	74.2	87	1	C89990	hypothetical prote	948	23	74.2	176	2	A99297	hypothetical prote
876	23	74.2	88	2	E71358	conserved hypotet	949	23	74.2	177	2	S65534	light-harvesting c
877	23	74.2	92	2	A87600	hypothetical prote	950	23	74.2	179	2	G95148	conserved hypotet
878	23	74.2	93	2	H90050	hypothetical prote	951	23	74.2	180	2	E98016	glutathione peroxi
879	23	74.2	95	2	AF1069	hypothetical prote	952	23	74.2	180	2	G97093	hypothetical prote
880	23	74.2	96	2	B41979	neuropeptide y pre	953	23	74.2	181	2	C69416	conserved hypotet
881	23	74.2	96	2	S49377	translation releas	954	23	74.2	182	2	G95979	hypothetical prote
882	23	74.2	96	2	B66701	protein P12A21.1 l	955	23	74.2	183	2	AE3613	conserved hypotet
883	23	74.2	102	2	S57536	hypothetical prote	956	23	74.2	184	2	S11242	amf10 protein - my
884	23	74.2	106	2	T46357	hypothetical prote	957	23	74.2	185	2	AG0717	probable dtrp-4-de
885	23	74.2	108	2	A13250	v1A/G regulated p	958	23	74.2	185	2	E95954	virulence protein
886	23	74.2	109	2	S12338	v1A/G regulated p	959	23	74.2	187	2	AC39185	hypothetical prote
887	23	74.2	109	2	C64402	hypothetical prote	960	23	74.2	188	2	AG2050	hypothetical prote
888	23	74.2	110	2	A11845	hypothetical prote	961	23	74.2	189	2	G89992	hypothetical prote
889	23	74.2	113	2	D83489	hypothetical prote	962	23	74.2	192	2	T15681	hypothetical prote
890	23	74.2	114	2	B97036	hypothetical prote	963	23	74.2	193	2	S53861	NADH2 dehydrogenas
891	23	74.2	120	2	T44411	probable metal-bin	964	23	74.2	195	2	E84272	hypothetical prote
892	23	74.2	122	1	R5SP14	ribosomal protein	965	23	74.2	197	2	AH2700	conserved hypotet
893	23	74.2	122	1	S24287	Ig heavy chain V r	966	23	74.2	198	2	G72471	hypothetical prote
894	23	74.2	123	1	R5NT14	ribosomal protein	967	23	74.2	200	2	AC0676	hypothetical prote
895	23	74.2	123	1	R5R214	ribosomal protein	968	23	74.2	202	2	T47571	H+-transporting tw
896	23	74.2	123	1	R5ZM14	ribosomal protein	969	23	74.2	203	2	D70111	GMP-binding protei
897	23	74.2	126	2	S72785	hypothetical prote	970	23	74.2	204	1	S71584	antioxidant Ahpc/
898	23	74.2	126	2	F70735	hypothetical prote	971	23	74.2	205	2	G97796	S-amino-6-(5)-phosp
899	23	74.2	126	2	H90097	hypothetical prote	972	23	74.2	206	2	E82287	hypothetical prote
900	23	74.2	127	2	T35336	probable isomerase	973	23	74.2	207	2	G84983	F7A19.24 protein -
901	23	74.2	128	2	T44877	hypothetical prote	974	23	74.2	207	2	T45600	hypothetical prote
902	23	74.2	129	2	F72425	2-amino-4-hydroxy-	975	23	74.2	207	2	T17343	hypothetical prote
903	23	74.2	132	2	B89834	ascidiac acid bios	976	23	74.2	208	2	B86275	hypothetical prote
904	23	74.2	135	2	UC7671	hypothetical prote	977	23	74.2	209	2	T26154	hypothetical prote
905	23	74.2	135	2	C93471	hypothetical prote	978	23	74.2	210	2		

979	23	74.2	211	2	E75415	hypothetical prote
980	23	74.2	211	2	F72239	hypothetical prote
981	23	74.2	212	2	T30700	hypothetical prote
982	23	74.2	212	2	T34052	hypothetical prote
983	23	74.2	213	2	S48402	hypothetical prote
984	23	74.2	213	2	T22991	hypothetical prote
985	23	74.2	215	2	C37390	transfer protein T
986	23	74.2	215	2	F96011	precortin-2 dehydr
987	23	74.2	216	1	C64081	L-fuculose phospho
988	23	74.2	218	1	F72214	hypothetical prote
989	23	74.2	221	2	S59075	splicing factor SR
990	23	74.2	222	2	AE3188	hypothetical prote
991	23	74.2	224	2	T10120	F420-dependent NAD
992	23	74.2	224	2	E72049	conserved hypother
993	23	74.2	224	2	F86575	CT691 hypothetical
994	23	74.2	226	2	T30615	hypothetical prote
995	23	74.2	230	2	E83554	probable pseudouri
996	23	74.2	231	2	G83809	MTA/SAH nucleosid
997	23	74.2	232	2	A69131	conserved hypother
998	23	74.2	232	2	S32963	hypothetical prote
999	23	74.2	233	2	T37376	27K structural pro
1000	23	74.2	234	2	T26363	hypothetical prote

ALIGNMENTS

RESULT 1
B90167
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 15-Jun-2001
C:Accession: B90167
R:She, O.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Aweyaz, M.J.; Chan
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi Ngoc, H.P.; Redder, R
arrett, R.A.; Ragan, M.A.; Sersen, C.W.; Van der Oost, J.
Submitted to Genbank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: B90167
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-281 <KUR>
A:Cross-references: GB:AE006641; NID:g13813391; PIDN:AAK40593.1; GSPDB:GN00155
C:Genetics:
A:Gene: SS00254
C:Superfamily: Archaeoglobus fulgidus conserved hypothetical protein AF0461

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 281;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
DB 160 VREEDL 165

RESULT 2
G75453
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: G75453
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M. Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 266, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75453
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-428 <WHI>

A:Cross-references: GB:AE001950; GB:AE000513; NID:g6458699; PIDN:AAF10548.1; PID:g645
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR0969
A:Map position: 1

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 428;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
DB 120 VREEDL 125

RESULT 3
JE0261
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) - human
C:Species: Homo sapiens (man)
C>Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 01-Mar-2002
C:Accession: JE0261
R:Uchimura, K.; Muramatsu, H.; Kaname, T.; Ogawa, H.; Yamakawa, T.; Fan, O.W.; Mitsuo
T. Biochem. 124, 670-678, 1998
A:Title: Human N-Acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis
A:Reference number: JE0261; MUID:98391845; PMID:9722682
A:Accession: JE0261
A:Molecule type: mRNA
A:Residues: 1-484 <UCH>
A:Cross-references: DDBJ:AB014679
C:Comment: This protein catalyzes the transfer of sulfur of sulfate from 3'-phosphoad
C:Superfamily: Chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 484;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
DB 377 VREEDL 382

RESULT 4
A64578
conserved hypothetical protein HP0465 - Helicobacter pylori (strain 26895)
C:Species: Helicobacter pylori
C>Date: 09-Aug-1997 #sequence_revision 09-Aug-1997 #text_change 08-Oct-1999
C:Accession: A64578
R:Tomb, J.F.; White, O.; Kerlavage, A.R.; Clayton, R.A.; Sutton, G.G.; Fleischmann, R
Peterson, S.; Loftus, B.; Richardson, D.; Dodson, R.; Khaliq, H.G.; Glodek, A.; McKe
son, U.D.; Kelley, J.M.; Cotton, M.D.; Weidman, J.M.; Fujii, C.; Bowman, C.; Watney,
Nature 388, 539-547, 1997
A:Authors: Wallin, E.; Hayes, W.S.; Borodovsky, M.; Karpk, P.D.; Smith, H.O.; Fraser,
A:Title: The complete genome sequence of the gastric pathogen Helicobacter pylori.
A:Reference number: A64520; MUID:97394467; PMID:9252185
A:Accession: A64578
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-631 <TOM>
A:Cross-references: GB:AE000561; GB:AE000511; NID:g2113564; PIDN:AAD07527.1; PID:g231

Query Match
Best Local Similarity 100.0%; Score 31; DB 2; Length 631;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
DB 182 VREEDL 187

RESULT 5

T30634
 hypothetical protein 32L - Molluscum contagiosum virus 1
 N:Alternate names: MC032L
 C:Species: Molluscum contagiosum virus 1
 C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jul-2000
 C:Accession: T30634
 R:Senkevich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
 Science 273, 813-816, 1996
 A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host re
 A:Reference number: 220876; MUID:56325459; PMID:8670425
 A:Accession: T30634
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-748 <SEN>
 A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:NA055160.1; PID:g1491975
 C:Genetics:
 A:Note: MC032L
 C:Superfamily: vaccinia virus hypothetical protein 49L

Query Match 100.0%; Score 31; DB 2; Length 748;
 Best Local Similarity 100.0%; Pred. No. 65;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 617 VREEDL 622

RESULT 6
 T21463
 hypothetical protein F28B1.2 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T21463
 R:Matthews, L.
 submitted to the EMBL Data Library, November 1996
 A:Reference number: Z19426
 A:Accession: T21463
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-973 <WIL>
 A:Cross-references: EMBL:Z81517; PIDN:CAB04209.1; GSPDB:GN00023; CESP:F28B1.2
 A:Experimental source: clone F28B1
 C:Genetics:
 A:Gene: CESP:F28B1.2
 A:Map position: 5
 A:Insertions: 70/3; 234/1; 267/3; 360/3; 600/3; 711/3; 926/3

Query Match 100.0%; Score 31; DB 2; Length 973;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 828 VREEDL 833

RESULT 7
 T30228
 polyketide synthase - Streptomyces hygroscopicus
 C:Species: Streptomyces hygroscopicus
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 17-Nov-2000
 C:Accession: T30228
 R:Paricio, J.F.; Molnar, I.; Schwecke, T.; Keenly, A.; Haydock, S.F.; Khaw, L.E.; Staun
 Gene 169, 9-16, 1996
 A:Title: Organisation of the biosynthetic gene cluster for rapamycin in Streptomyces hyg
 A:Reference number: Z20782; MUID:96186896; PMID:8635756
 A:Accession: T30228
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-6260 <APA>
 A:Cross-references: EMBL:X86780; NID:g987088; PID:g987102; PIDN:CAA60462.1
 C:Genetics:

A:Gene: rapC
 C:Superfamily: 3-oxoacyl-[acyl-carrier-protein] synthase I homology; acyl carrier pro
 C:Keywords: carrier protein
 F:53-451/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS1>
 F:1536-1607/Domain: acyl carrier protein homology <ACP1>
 F:1651-2042/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>
 F:3123-3194/Domain: acyl carrier protein homology <ACP2>
 F:3238-3632/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS3>
 F:3725-3999/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F:5114-5185/Domain: acyl carrier protein homology <ACP3>
 F:5242-5638/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS4>
 F:5759-6044/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT2>
 F:6135-6206/Domain: acyl carrier protein homology <ACP4>

Query Match 100.0%; Score 31; DB 2; Length 6260;
 Best Local Similarity 100.0%; Pred. No. 6.7e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 4365 VREEDL 4370

RESULT 8
 JC7628
 glucoside 3-dehydrogenase (EC 1.1.99.13) - Halomonas sp. alpha-15
 C:Species: Halomonas sp. alpha-15
 C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
 C:Accession: JC7628; PC7123
 R:Kojima, K.; Tsugawa, W.; Sode, K.
 Biochem. Biophys. Res. Commun. 282, 21-27, 2001
 A:Title: Cloning and expression of glucose 3-dehydrogenase from Halomonas sp. alpha-1
 A:Reference number: JC7628; MUID:21164693; PMID:11263965
 A:Accession: JC7628
 A:Molecule type: DNA
 A:Residues: 1-562 <KOJ>
 A:Accession: PC7123
 A:Molecule type: protein
 A:Residues: 2-11 <KO2>
 C:Comment: This enzyme, as an oligomeric enzyme composed of catalytic and electron tr
 and disaccharides.
 C:Genetics:
 A:Gene: q3dh
 C:Keywords: oxidoreductase

Query Match 96.8%; Score 30; DB 2; Length 562;
 Best Local Similarity 83.3%; Pred. No. 80;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 142 VREEDL 147

RESULT 9
 HB7451
 oxidoreductase, GMC family C01634 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
 C:Accession: HB7451
 R:Nielemann, W.C.; Feldlynn, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko
 n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of Caulobacter crescentus.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: HB7451
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-579 <STO>
 A:Cross-references: GB:AE005673; NID:g13423038; PIDN:AKK23612.1; GSPDB:GN00148
 C:Genetics:
 A:Gene: C01634

Query Match 96.8%; Score 30; DB 2; Length 579;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 153 IREYDL 158

RESULT 10
B71440
hypothetical protein d14605c - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
A:Variety: Columbia
C:Date: 03-Aug-1998 #sequence_revision 03-Aug-1998 #text_change 20-Jun-2000
C:Accession: B71440
R:Byan, M.; Hancock, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk
P.; Medler, H.; Medler, E.; Wambutt, R.; Woltzenegger, T.; Pohl, T.M.; Terry, N.; Giel
avanagh, T.; Hempel, S.; Kotzer, P.; Ertan, K.D.; Rieger, M.; Schaeffer, M.; Funk, B.
Nature 391, 485-488, 1998
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans
C.; Chalmatis, N.
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis tha
A:Reference number: A71400; MUID:98121113; PMID:9461215
A:Accession: B71440
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1322 <ENV>
A:Cross-references: GB:297342; NID:92245031; PID:92245072
C:Genetics:
A:Map position: 4COP9-4G3845
C:Superfamily: Arabidopsis thaliana hypothetical protein d14605c

Query Match 96.8%; Score 30; DB 2; Length 1322;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 21 IREYDL 26

RESULT 11
H87385
hypothetical protein CC1100 [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: H87385
R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Land, M.I.; DeBoy, R.T.; Dodson, R.O.; Durkin, A.S.; Gwinn, M.L.; Hart, D.H.; Kolot
n, J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87385
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-734 <STO>
A:Cross-references: GB:AE005673; NID:913422406; PIDN:AAK23084.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC1100

Query Match 93.5%; Score 29; DB 2; Length 234;
Best Local Similarity 83.3%; Pred. No. 52;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 146 VREYDL 151

RESULT 12
F96020
probable iron ABC transporter periplasmic solute-binding protein precursor SMD20723 [
C:Species: Sinorhizobium meliloti
C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: F96020
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A:Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N₂-fixing e
A:Reference number: A95842; MUID:21396508; PMID:11481431
A:Accession: F96020
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-355 <XUR>
A:Cross-references: GB:AL591985; PIDN:CAC49830.1; PID:915141318; GSPDB:GN00167
A:Experimental source: strain 1021, megaplasmid pSymB
R:Gallagher, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Ruhl
pel, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMD20723
A:Genome: plasmid

Query Match 93.5%; Score 29; DB 2; Length 355;
Best Local Similarity 83.3%; Pred. No. 82;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 63 VREYDM 68

RESULT 13
T08203
polygalacturonase (EC 3.2.1.15) precursor - muskmelon
C:Species: Cucumis melo (muskmelon)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-May-2000
C:Accession: T08203
R:Hadielid, K.A.; Rose, J.K.; Yaver, D.S.; Berka, R.M.; Bennett, A.B.
Plant Physiol. 117, 363-373, 1998
A:Title: Polygalacturonase gene expression in ripe melon fruit supports a role for po
A:Reference number: Z16403; MUID:98289082; PMID:9825689
A:Accession: T08203
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-393 <HAD>
A:Cross-references: EMBL:AF062465; NID:93320457; PID:93320458
C:Genetics:
A:Note: MP1
C:Superfamily: polygalacturonase
C:Keywords: glycosidase; hydrolase

Query Match 93.5%; Score 29; DB 2; Length 393;
Best Local Similarity 83.3%; Pred. No. 92;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
:|||||
DB 329 VREYDI 334

RESULT 14
D84871
probable polygalacturonase [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001
C:Accession: D84871

R:lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.: Koo, H.; Morfat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*.

A:Reference number: A84420; MUID:20083487; PMID:1061197

A:Accession: D84871

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1,405 <STC>

A:Cross-references: GB:AE002093; NID:g2281084; PIDN:AAB64020.1; GSPDB:GN00139

C:Genetics:

A:Gene: AT2943860

A:Map position: 2

C:Superfamily: polygalacturonase

Query Match 93.5%; Score 29; DB 2; Length 405;
Best Local Similarity 83.3%; Pred. No. 95;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 335 VREYDL 340

RESULT 15
G64446
ATP-dependent 26S proteasome regulatory subunit 4 homolog - *Methanococcus jannaschii*
C:Species: *Methanococcus jannaschii*
C>Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 19-Jan-2001
C:Accession: G64446
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, J.; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.; Ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A. Science 273, 1058-1073, 1996
A:Authors: Kaine, B.P.; Borodovsky, M.; Klein, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*
A:Reference number: A64300; MUID:96337999; PMID:8688087
A:Accession: G64446
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1,430 <BUL>
A:Cross-references: GB:U67559; GB:L77117; NID:g1591798; PIDN:AAB99179.1; PID:g1591803; C:Genetics:
A:Map position: REV1116282-1114990
C:Superfamily: ATP-dependent 26S proteasome: FtsH/SEC18/CDC48-type ATP-binding domain hc
C:Keywords: ATP; nucleotide binding; P-loop
F:184-394/Domain: FtsH/SEC18/CDC48-type ATP-binding domain homology <VAMP>
F:211-218/Region: nucleotide-binding motif A (P-loop)

Query Match 93.5%; Score 29; DB 2; Length 430;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 167 VREYDL 172

RESULT 16
T39225
MAP kinase kinase kinase - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Jan-2000
C:Accession: T39225
R:Churcher, C.M.; Gentile, S.; Barrell, B.G.; Rajandream, M.A.; Wood, V. submitted to the EMBL Data Library, August 1997
A:Reference number: 221837
A:Accession: T39225
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1401 <CHU>
A:Cross-references: EMBL:298763; PIDN:CA11500.1; GSPDB:GN00066; SPDB:SPAC961.02

A:Experimental source: strain 972h-; cosmid c961
C:Genetics:
A:Gene: SPDB:SPAC961.02
A:Map position: 1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase hom

Query Match 93.5%; Score 29; DB 2; Length 1401;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 1194 VREYDL 1199

RESULT 17
D71350
probable DNA-directed RNA polymerase, beta' subunit - *Syphilis spirochete*
C:Species: *Treponema pallidum* subsp. *pallidum* (*Syphilis spirochete*)
C>Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 18-Jun-1999
C:Accession: D71350
R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Ison, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uitterback, T.; G they, L.; Weidman, J.; Smith, H.O.; Venter, J.C. Science 281, 375-388, 1998
A:Title: Complete genome sequence of *Treponema pallidum*, the *syphilis spirochete*.
A:Reference number: A71250; MUID:9832770; PMID:9665876
A:Accession: D71350
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1,1416 <COL>
A:Cross-references: GB:AE001205; GB:AE000520; NID:g3322501; PIDN:AAC65230.1; PID:g3332
A:Experimental source: strain Nichols
A:Gene: TP0242
C:Superfamily: *Escherichia coli* DNA-directed RNA polymerase beta' chain

Query Match 93.5%; Score 29; DB 2; Length 1416;
Best Local Similarity 83.3%; Pred. No. 3.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||
DB 1052 VREYDL 1057

RESULT 18
B69441
conserved hypothetical protein AF1531 - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: B69441
R:Kleck, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod . Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L. Nature 390, 364-370, 1997
A:Authors: Uitterback, T.; Cotton, M.D.; Spriggs, T.; Artlich, P.; Kaine, B.P.; Sykes, Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343; PMID:9385475
A:Accession: B69441
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-195 <KLE>
A:Cross-references: GB:AE000997; GB:AE000782; NID:92689320; PIDN:AAB89717.1; PID:g264
C:Superfamily: *Methanococcus jannaschii* conserved hypothetical protein Mj0038

Query Match 90.3%; Score 28; DB 1; Length 195;
Best Local Similarity 83.3%; Pred. No. 72;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||

Db 87 LRYEDL 92

RESULT 19

C69608 regulatory protein [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C:Accession: C69608

R:Botolin, A.; Winkler, P.; Mauger, S.; Jallou, O.; Malarme, K.; Weissenbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ss

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: C69608

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-233 <STO>

A:Cross-references: GB:AE005176; PID:g12724820; PIDN:AAK05893.1; GSPDB:GN00146

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: ysfD

Query Match 90.3%; Score 28; DB 2; Length 233;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 146 IRYEDI 151

RESULT 20

T42928 Immediate-early protein - ateline herpesvirus 3 (strain 73)

C:Species: ateline herpesvirus 3

A:Variety: strain 73

C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 18-Feb-2000

C:Accession: T42928

R:Albrecht, J.C.; Fleckenstein, B.

Submitted to the EMBL Data Library, August 1998

A:Description: Primary structure of the herpesvirus ateles genome.

A:Reference number: 222274

A:Accession: T42928

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-273 <ALB>

A:Cross-references: EMBL:AF083424; PIDN:AA095536.1

A:Experimental source: strain 73

C:Superfamily: saimiri herpesvirus immediate-early protein 1

Query Match 90.3%; Score 28; DB 2; Length 273;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 208 VRYODL 213

RESULT 21

C69608 chitosanase csn - Bacillus subtilis

C:Species: Bacillus subtilis

C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999

C:Accession: C69608

R:Kunst, F.; Ogatawa, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter

C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

bach, J.; Hartwood, C.R.; Henuit, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.

Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,

A:Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel

Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portete

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scani

A:Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekituchi, J.; Sekowska, A.; Se

akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Tempere, P.; Tognoni, A.; Tosato, V.; Uchiya

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida

A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtili

A:Reference number: A69580; MUID:98044033; PMID:9384377

A:Accession: C69608

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-277 <KUN>

A:Cross-references: GB:Z99117; GB:AI009126; NID:G2634966; PIDN:CB14630.1; PID:ell839

A:Experimental source: strain 168

C:Genetics:

A:Gene: csn

Query Match 90.3%; Score 28; DB 2; Length 277;
Best Local Similarity 83.3%; Pred. No. 1.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 222 VRYDDL 227

RESULT 22

A83634 hypothetical protein PA0100 [imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: A83634

R:Scovier, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L

.. Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337; PMID:10984043

A:Accession: A83634

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-306 <STO>

A:Cross-references: GB:AE004448; GB:AE004091; NID:g9945917; PIDN:AA03490.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA0100

Query Match 90.3%; Score 28; DB 2; Length 306;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||:|
Db 16 VRYODL 21

RESULT 23

A84523 probable steroid sulfotransferase [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 16-Feb-2001

C:Accession: A84523

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujita, C.Y

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,

euss, D.; Nieman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:1061197

A:Accession: A84523

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-333 <STO>

A:Cross-references: GB:AE002093; NID:g3650034; PIDN:AA061289.1; GSPDB:GN00139

C:Genetics:
A:Gene: At2g14920
A:Map position: 2
C:Superfamily: alcohol sulfotransferase

Query Match 90.3%; Score 28; DB 2; Length 333;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 232 MRYEDL 237

RESULT 24

C38453
anaerobic sulfite reduction protein C - *Salmonella typhimurium*
C:Species: *Salmonella typhimurium*
C:Date: 12-Jul-1991 #sequence_revision 12-Jul-1991 #text_change 17-Mar-2000
C:Accession: C38453
R:Huang, C.J.; Barrett, E.L.
J. Bacteriol. 173, 1544-1553, 1991
A:Title: Sequence analysis and expression of the *Salmonella typhimurium* asr operon encoded
A:Reference number: A38453; M01D:91139599; PMID:1704886
A:Accession: C38453
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <HUA>
A:Cross-references: GB:M57706; NID:g153881; PIDN:AA99277.1; PID:g153884
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
C:Keywords: 4Fe-4S; iron-sulfur protein; metalloprotein
F:173-230/Domain: ferredoxin 2[4Fe-4S] homology <FER>
F:180,183,186,222/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted
F:190,212,215,218/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 90.3%; Score 28; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 69 IRYEDL 74

RESULT 25

AE0825
anaerobic sulfite reductase chain C (EC 1.8.1.-) [imported] - *Salmonella enterica* subsp.
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh
A:Note: this species has also been called *Salmonella typhi*
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 27-Nov-2001
C:Accession: AE0825
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, R.; Davies, R.; Groth, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Mout, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
A:Reference number: AB0502; PMID:11677608
A:Accession: AE0825
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-337 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD02753.1; PID:g16503764; GSPDB:GN00176
C:Genetics:
A:Gene: asrC
C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] h
C:Keywords: oxidoreductase

Query Match 90.3%; Score 28; DB 2; Length 337;
Best Local Similarity 66.7%; Pred. No. 1.3e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

DB 69 IRYEDL 74

RESULT 26

T29003
hypothetical protein ZC513.9 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T29003
R:Wu, X.; Le, T.T.
submitted to the EMBL Data Library, April 1996
A:Description: The sequence of C. elegans cosmid ZC513.
A:Reference number: Z20551
A:Accession: T29003
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-346 <WUX>
A:Cross-references: EMBL:U53155; PIDN:AAC48271.1; GSPDB:GN00023; CESP:ZC513.9
C:Genetics:
A:Gene: CESP:ZC513.9
A:Map position: 5
A:Introns: 62/3; 117/2; 252/3; 312/1

Query Match 90.3%; Score 28; DB 2; Length 346;
Best Local Similarity 83.3%; Pred. No. 1.3e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 79 LRYEDL 84

RESULT 27

G70729
hypothetical protein RV2267c - *Mycobacterium tuberculosis* (strain H37RV)
C:Species: *Mycobacterium tuberculosis*
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: G70729
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtz, R.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno
A:Reference number: A70500; M01D:98295987; PMID:9634230
A:Accession: G70729
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-388 <COL>
A:Cross-references: GB:Z77163; GB:AL123456; NID:g3261610; PIDN:CAB00968.1; PID:e25507
C:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV2267c

Query Match 90.3%; Score 28; DB 2; Length 386;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 302 LRYEDL 307

RESULT 28

AH2447
molybdopterin biosynthesis protein [imported] - *Nostoc* sp. (strain PCC 7120)
C:Species: *Nostoc* sp.
A:Note: *Nostoc* sp. strain PCC 7120 is a synonym of *Anabaena* sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 30-Jun-2002
C:Accession: AH2447
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimpō, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Res. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium *Anaerostipes*
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: AB2447
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-436 <KUR>
A:Cross-references: GB:BA000019; PIDN:BA576835.1; PID:g17134274; GSPDB:GN00179
A:Experimental source: strain FCC 7120
A:Genetics: moea
C:Superfamily: molybdenum cofactor biosynthesis protein moea-2

Query Match 90.3%; Score 28; DB 2; Length 456;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 84 VREEDV 89

RESULT 29

C75180
tld related protein PAB1955 - Pyrococcus abyssi (strain Orsay)
C:Species: Pyrococcus abyssi
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C75180
R:anonymus; Genoscope
submitted to the EMBL Data Library, July 1999
A:Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome stru
A:Reference number: A75001
A:Accession: C75180
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-458 <KAM>
A:Cross-references: GB:A0248284; GB:A1096836; NID:95457730; PIDN:CB49522.1; PID:g545803
A:Experimental source: strain Orsay
C:Genetics:
A:Gene: PAB1955
C:Superfamily: Escherichia coli tld protein

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 25 LREEDL 30

RESULT 30

A57397
chondroitin 6-sulfotransferase (EC 2.8.2.17) precursor - chicken
C:Species: Gallus gallus (chicken)
C>Date: 23-Feb-1996 #sequence_revision 23-Feb-1996 #text_change 20-Jun-2000
C:Accession: A57397
R:Fukuta, M.; Uchimura, K.; Nakashima, K.; Kato, M.; Kimata, K.; Shinomura, T.; Habuchi,
J. Biol. Chem. 270, 18575-18580, 1995
A:Title: Molecular cloning and expression of chick chondrocyte chondroitin 6-sulfotransf
A:Reference number: A57397; MUID:95355490; PMID:7629189
A:Accession: A57397
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-458 <FUK>
A:Cross-references: GB:D49915; NID:9711262; PIDN:BA08655.1; PID:g9711263
C:Superfamily: chondroitin 6-sulfotransferase
C:Keywords: sulfotransferase

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 348 VREEDV 353

RESULT 31

JC7350
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - mouse
C:Species: Mus musculus (house mouse)
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: JC7350
R:Uchimura, K.; Fasaki, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa,
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of
A:Reference number: JC7350; MUID:20374462; PMID:10913333
A:Accession: JC7350
A:Molecule type: mRNA
A:Residues: 1-484 <UCH>
A:Cross-references: DDBJ:AB040710
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosyl
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 90.3%; Score 28; DB 2; Length 484;
Best Local Similarity 83.3%; Pred. No. 1.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 371 LREEDL 376

RESULT 32

JC7351
N-acetylglucosamine-6-O-sulfotransferase (EC 2.8.2.-) [validated] - human
C:Species: Homo sapiens (man)
C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 31-Dec-2000
C:Accession: JC7351
R:Uchimura, K.; Fasaki, F.; Kadomatsu, K.; Matsukawa, T.; Yamakawa, T.; Kurosawa,
Biochem. Biophys. Res. Commun. 274, 291-296, 2000
A:Title: Diversity of N-acetylglucosamine-6-O-sulfotransferases: Molecular cloning of
A:Reference number: JC7350; MUID:20374462; PMID:10913333
A:Accession: JC7351
A:Molecule type: mRNA
A:Residues: 1-486 <UCH>
A:Cross-references: DDBJ:AB040711
C:Comment: This enzyme, having a type II transmembrane topology and N-linked glycosyl
sulfate linkage formation. This enzyme is involved in synthesis of L-selectin ligand
C:Keywords: glycolysis; sulfate transport; sulfotransferase; transmembrane protein

Query Match 90.3%; Score 28; DB 2; Length 486;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 373 LREEDL 378

RESULT 33

F90169
hypothetical protein tgra [imported] - Sulfolobus solfataricus
C:Species: Sulfolobus solfataricus
C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 24-May-2001
C:Accession: F90169
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Ch
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder
arlett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A:Description: Sulfolobus solfataricus complete genome.
A:Reference number: A99139
A:Accession: F90169

Query Match 90.3%; Score 28; DB 2; Length 486;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-503 <RUR>
A:Cross-references: GB:AE006641; NID:g13813415; PIDN:AAK40613.1; GSPDB:GN00155
C:Genetics:
A:Gene: tga

Query Match 90.3%; Score 28; DB 2; Length 503;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 6 VREEDL 11

RESULT 34

E89804
alkyl hydroperoxide reductase subunit F [imported] - *Staphylococcus aureus* (strain N315)
C:Species: *Staphylococcus aureus*
C:Date: 10-May-2001 #sequence-revision 10-May-2001 #text-change 22-Oct-2001
C:Accession: E89804
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiratsugu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89756; MUID:21311952; PMID:11418146
A:Accession: E89804
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-507 <KUR>
A:Cross-references: GB:BA000018; PID:g13700294; PIDN:BAK41592.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: amp
C:Superfamily: NADH oxidase (hydrogen peroxide-forming); thioredoxin reductase homology

Query Match 90.3%; Score 28; DB 2; Length 507;
Best Local Similarity 66.7%; Pred. No. 2.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 416 IREEDM 421

RESULT 35

T45641
beta-D-glucan exohydrolase-like protein - *Arabidopsis thaliana*
N:Alternate names: protein F13112.100
C:Species: *Arabidopsis thaliana* (mouse-ear cress)
C:Date: 04-Feb-2000 #sequence-revision 04-Feb-2000 #text-change 21-Jul-2000
C:Accession: T45641
R:Choinsne, N.; Robert, C.; Brotier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sa

submitted to the Protein Sequence Database, November 1999
A:Reference number: Z23010
A:Accession: T45641
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-612 <CHO>
A:Cross-references: EMBL:AL133292
A:Experimental source: cultivar Columbia; BAC clone F13112
C:Genetics:

A:Map position: 3
A:Introns: 56/1; 124/2; 154/3; 203/2; 284/3; 319/3; 380/3; 446/1
A:Note: F13112.100
C:Superfamily: beta-glucosidase

Query Match 90.3%; Score 28; DB 2; Length 612;
Best Local Similarity 83.3%; Pred. No. 2.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 230 IREEDL 235

RESULT 36

B91271
probable periplasmic binding protein [imported] - *Escherichia coli* (strain O157:H7, s
C:Species: *Escherichia coli*
C:Date: 18-Jul-2001 #sequence-revision 18-Jul-2001 #text-change 18-Jul-2001
C:Accession: B91271
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g
A:Reference number: A99629; MUID:21156231; PMID:11258796
A:Accession: B91271
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1107 <HAV>
A:Cross-references: GB:BA000007; PIDN:BA38561.1; PID:g13364615; GSPDB:GN00154
A:Experimental source: strain O157:H7, substrain RMD 0509952
C:Genetics:
A:Gene: Ecs5138

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 348 IREEDL 353

RESULT 37

B86112
probable periplasmic binding protein yJep [imported] - *Escherichia coli* (strain O157:
C:Species: *Escherichia coli*
C:Date: 16-Feb-2001 #sequence-revision 16-Feb-2001 #text-change 14-Sep-2001
C:Accession: B86112
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apoda
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: B86112
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1107 <STO>
A:Cross-references: GB:AE005174; NID:g12519144; PIDN:AAG59358.1; GSPDB:GN00145; OMGP:
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: yJep

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 348 IREEDL 353

RESULT 38

E65226
hypothetical 123.8 kD protein in genX-psd intergenic region - *Escherichia coli* (strai
N:Alternate names: hypothetical protein F1107
C:Species: *Escherichia coli*
C:Date: 12-Sep-1997 #sequence-revision 17-Sep-1997 #text-change 01-Mar-2002
C:Accession: E65226; S56387
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997

A:Title: The complete genome sequence of *Escherichia coli* K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E65226
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1107 <BLAT>
A:Cross-references: GB:AE000489; GB:U00096; NID:g2367354; PIDN:AAC77119.1; PID:g2367355
A:Experimental source: strain K-12, substrain MG1655
A:Buriald, V.; Plunkett III, G.; Sofia, H.J.; Daniels, D.L.; Blattner, F.R.
Nucleic Acids Res. 23, 2105-2119, 1995
A:Title: Analysis of the *Escherichia coli* genome VI: DNA sequence of the region from 92.
A:Reference number: S56314; MUID:95333636; PMID:7610040
A:Accession: S56387
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-746, 'X', 748-1014, 'A', 1016-1107 <BUR>
A:Cross-references: EMBL:U14003; NID:g1263172; PIDN:AA97058.1; PID:g537003
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1994
C:Genetics:
A:Start codon: GTG
A:Start codon: GTG

Query Match 90.3%; Score 28; DB 2; Length 1107;
Best Local Similarity 83.3%; Pred. No. 4.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 348 LREEDL 353

RESULT 39
ZIN2P3
genome polypeptide - parainfluenza virus type 3
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C:Species: parainfluenza virus type 3
A:Note: host Homo sapiens (man)
C:Date: 31-Dec-1989 #sequence_revision 31-Dec-1989 #text_change 29-May-1998
C:Accession: A29246
R:Galinski, M.S.; Mink, M.A.; Pons, M.W.
Virology 165, 499-510, 1988
A:Title: Molecular cloning and sequence analysis of the human parainfluenza 3 virus gene
A:Reference number: A29246; MUID:88306242; PMID:2841798
A:Accession: A29246
A:Molecule type: genomic RNA
A:Residues: 1-2233 <GAL>
A:Cross-references: GB:M21649
C:Genetics:
A:Gene: L
C:Superfamily: parainfluenza virus RNA-directed RNA polymerase
C:Keywords: ATP; nucleotidyltransferase

Query Match 90.3%; Score 28; DB 1; Length 2233;
Best Local Similarity 66.7%; Pred. No. 1e+03;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 1126 LREEDL 1131

RESULT 40
H83960
hypothetical protein BH2488 [Imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: H83960
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and
A:Reference number: A83650; MUID:2051582; PMID:11058132
A:Accession: H83960
A:Status: preliminary

A:Molecule type: DNA
A:Residues: 1-73 <STO>
A:Cross-references: GB:AP001515; GB:BA000004; NID:g10174886; PIDN:BA06207.1; GSPDB:G
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH2488

Query Match 87.1%; Score 27; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 39 RYEDL 43

RESULT 41
F72649
hypothetical protein APE0628 - *Aeropyrum pernix* (strain K1)
C:Species: *Aeropyrum pernix*
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jun-2000
C:Accession: F72649
R:Kawabuchi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-no, K.; Ta
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic *Crenarchaeon*, *Aero*
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72649
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-108 <KAW>
A:Cross-references: DDBJ:AP000060; NID:g5104188; PIDN:BA79598.1; PID:d1043384; PID:g
C:Genetics:
A:Gene: APE0628
C:Superfamily: *Aeropyrum pernix* hypothetical protein APE0628

Query Match 87.1%; Score 27; DB 2; Length 108;
Best Local Similarity 83.3%; Pred. No. 63;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 16 VREEDL 21

RESULT 42
H69433
response regulator homolog - *Archaeoglobus fulgidus*
C:Species: *Archaeoglobus fulgidus*
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 20-Aug-1999
C:Accession: H69433
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod
gloek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weisman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Ullrich, T.; Cotton, M.D.; Spriggs, T.; Artach, P.; Kaine, B.P.; Sykes,
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: H69433
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-133 <KLE>
A:Cross-references: GB:AE001001; GB:AE000782; NID:g2689324; PIDN:AAB9775.1; PID:g264
C:Superfamily: chemotaxis chey protein; response regulator homolog
C:Keywords: phosphoprotein
F:4.11/Domain: response regulator homolog <RRH>
F:54/Binding site: phosphate (Asp) (covalent) #status predicted

Query Match 87.1%; Score 27; DB 2; Length 133;
Best Local Similarity 83.3%; Pred. No. 80;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
11111
Db 108 VREEDL 113

RESULT 43

A23571
myosin LC1 catalytic light chain - Japanese flying squid
C:Species: Todarodes pacificus (Japanese flying squid)
C>Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 03-May-1996
C:Accession: A23571
R:Matanabe, B.; Maite, T.; Komno, K.; Matsuda, G.
Biol. Chem. Hoppe-Seyler 367, 1025-1032, 1986
A:Title: Amino-acid sequence of LC-1 light chain of squid mantle muscle myosin.
A:Reference number: A25571; MUID:87076038; PMID:3790251
A:Accession: A25571
A:Molecule type: protein
A:Residues: 1-160 <KUR>
C:Superfamily: calmodulin; calmodulin repeat homology
C:Keywords: calcium binding; EF hand; muscle
F:7-41/Domain: calmodulin repeat homology <EF1>
F:45-76/Domain: calmodulin repeat homology <EF2>
F:83-115/Domain: calmodulin repeat homology <EF3>
F:121-153/Domain: calmodulin repeat homology <EF4>

Query Match 87.1%; Score 27; DB 2; Length 160;
Best Local Similarity 66.7%; Pred. No. 98;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
11111
Db 141 IREEDL 146

RESULT 44

C95315
hypothetical protein SMA0791 [imported] - Sinorhizobium meliloti (strain 1021) magaplasma
C:Species: Sinorhizobium meliloti
C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
C:Accession: C95315
R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows,
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A:Reference number: A95262; MUID:21396509; PMID:11481432
A:Accession: C95315
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-200 <KUR>
A:Cross-References: GB:AE006469; PIDN:AAK65085.1; PID:g14523520; GSPDB:GN00465
A:Experimental source: strain 1021, megaplasmid pSymA
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A:Reference number: A96039; MUID:21368234; PMID:11474104
A:Contents: annotation
C:Genetics:
A:Gene: SMA0791
A:Genome: plasmid

Query Match 87.1%; Score 27; DB 2; Length 200;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
11111
Db 183 RYEDL 187

RESULT 45

C95194
MutT/nudix family protein [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C>Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: C95194
R:Rettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; H
on, J.D.; Umayam, L.A.; White, O.; Salberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapf
nson, T.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morris
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
A:Reference number: A95000; MUID:21357209; PMID:11463916
A:Accession: C95194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <KUR>
A:Cross-References: GB:AE005672; PIDN:AAK5748.1; PID:g14973161; GSPDB:GN00164; TIGR:
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP1669

Query Match 87.1%; Score 27; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
11111
Db 32 RYEDL 36

RESULT 46

H98060
mutator protein [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: H98060
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.
e, R.; LeBlanc, D.J.; Lee, L.N.; Letkowitz, E.J.; Lu, J.; Matsushima, P.; McAnren, S.
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: H98060
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-203 <KUR>
A:Cross-References: GB:AE007317; PIDN:AAU00317.1; PID:g15459175; GSPDB:GN00174
C:Genetics:
A:Gene: mutT

Query Match 87.1%; Score 27; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
11111
Db 32 RYEDL 36

RESULT 47

G70340
Mg(2+) transport ATPase - Aquifex aeolicus
C:Species: Aquifex aeolicus
C>Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000
C:Accession: G70340
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.;
V.
Nature 392, 353-358, 1998
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A:Reference number: A70300; MUID:98196666; PMID:9537320

A:Accession: G70340
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-225 <AOF>
A:Cross-references: GB:AE000690; NID:92983100; PIDN:AAC06713.1; PID:92983108; GB:AE00069
A:Experimental source: strain VFS
A:Genetics:
A:Gene: mgc
C:Superfamily: Mg2+-transporting ATPase

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 225;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 VRYED 6
|||||
Db 202 VRYED 206

RESULT 48
S28679
glycoprotein 30 - phage SP01
C:Species: phage SP01
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
C:Accession: S28679
R:Scarlato, V.; Sayre, M.H.
Gene 114, 115-119, 1992
A:Title: Sequence of the bacteriophage SP01 gene 30.
A:Reference number: S28678; MUID:92267370; PMID:1587473
A:Accession: S28679
A:Molecule type: DNA
A:Residues: 1-225 <SCA>
A:Cross-references: EMBL:M82842; NID:9216115; PIDN:AAA32597.1; PID:9216117
C:Genetics:
A:Start codon: TTG
C:Superfamily: phage SP01 glycoprotein 30
C:Keywords: glycoprotein

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 225;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 32 VRYED 36

RESULT 49
C90026
hypothetical protein narQ [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: C90026
R:Kurita, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: C90026
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-265 <KUR>
A:Cross-references: GB:BA000018; PID:913702080; PIDN:BA843372.1; GSPDB:GN00149
A:Experimental source: strain N315
A:Genetics:
C:Superfamily: Escherichia coli fdhD protein

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 265;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 12 VRYED 16

RESULT 50
S07014
hypothetical protein 273 (orf 5' region) - Escherichia coli plasmid F
C:Species: Escherichia coli
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 21-Jul-2000
C:Accession: S07014
R:Loch, S.; Cram, D.; Skurray, R.
Mol. Gen. Genet. 219, 177-186, 1989
A:Title: Nucleotide sequence of the leading region adjacent to the origin of transfer
A:Reference number: S07013; MUID:90136505; PMID:2693941
A:Accession: S07014
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-273 <LOH>
A:Cross-references: GB:AF106329; GB:M97768; GB:M21120; GB:X17189; NID:95702165; PIDN:
A:Genetics:
A:Genome: plasmid
C:Superfamily: Escherichia hypothetical protein yafZ

Query Match
Best Local Similarity 87.1%; Score 27; DB 2; Length 273;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 189 VRYED 193

Search completed: February 20, 2003, 13:34:39
Job time : 26.6 secs

107	25	80.6	346	1	YQCL_BACSU	P45946	bacillus su	180	24	77.4	235	1	VG53_BPMD2	O64243	mycobacteri
108	25	80.6	350	1	Y535_METUA	Q57975	methanococc	181	24	77.4	245	1	CRS3_HUMAN	P54108	homo sapien
109	25	80.6	366	1	YVAF_BACSU	P37518	bacillus su	182	24	77.4	245	1	PSBP_CHURE	P14711	chlamydomon
110	25	80.6	375	1	PCF_TACTR	P21902	tachypleus	183	24	77.4	257	1	NUDC_ECO57	Q8x6x7	escherichia
111	25	80.6	409	1	EFTU_EUGGR	P02991	euglena gra	184	24	77.4	257	1	NUDC_ECOLI	P32664	escherichia
112	25	80.6	414	1	PYRC_THEAC	Q9h1m0	thermoplasm	185	24	77.4	257	1	NUDC_SALTI	Q8x328	salmoneilla
113	25	80.6	436	1	THCB_RHOER	P43492	rhodococcus	186	24	77.4	257	1	NUDC_SALTY	Q91915	salmoneilla
114	25	80.6	452	1	SELA_AQUAE	C67140	aquifex aeo	187	24	77.4	258	1	ELAP_ECOLI	P43530	escherichia
115	25	80.6	486	1	DNAB_HELPJ	Q92jms	helicobacte	188	24	77.4	258	1	PSBP_ECOLI	P06717	escherichia
116	25	80.6	488	1	DNAB_HELPY	O25916	helicobacte	189	24	77.4	258	1	PSBP_LYCES	P25795	lycopersico
117	25	80.6	491	1	GLNA_ARCFU	O29313	archaeoglob	190	24	77.4	259	1	PSBP_WHEAT	O00434	tritlicum ae
118	25	80.6	520	1	GAL3_YEAST	P13045	saccharomyc	191	24	77.4	259	1	PSBP_PEA	P16059	plum sativ
119	25	80.6	532	1	YMOO_YEAST	O04488	saccharomyc	192	24	77.4	259	1	Y128_CHLPP	P11594	strapsis alb
120	25	80.6	547	1	TEPH_MOUSE	Q32780	homo sapien	193	24	77.4	260	1	PSBP_SINAL	P93566	solanum tub
121	25	80.6	548	1	TEPH_MOUSE	Q32780	homo sapien	194	24	77.4	260	1	PSBP_SORU	O9s198	cucumis sat
122	25	80.6	573	1	GIP1_YEAST	P38229	saccharomyc	195	24	77.4	263	1	PSBP_CUGSA	O9s198	cucumis sat
123	25	80.6	585	1	YHD9_YEAST	P38732	saccharomyc	196	24	77.4	264	1	PSBP_PRIAG	O45080	trillilaria
124	25	80.6	605	1	VE1_HPV09	Q05111	human papil	197	24	77.4	265	1	PSBP_NARPS	Q04040	narcissus p
125	25	80.6	609	1	VE1_HPV17	O02512	human papil	198	24	77.4	265	1	SANT_PLAFI	P04927	plasmidum
126	25	80.6	616	1	ORC3_YEAST	P54750	saccharomyc	199	24	77.4	267	1	PSBP_SPIOL	P12302	spniacia ol
127	25	80.6	620	1	XG18_YEAST	P53207	saccharomyc	200	24	77.4	268	1	VBGF_PSEPU	P43037	pseudomonas
128	25	80.6	626	1	MK33_HUMAN	Q99799	saccharomyc	201	24	77.4	271	1	KKAI_ECOLI	P00551	escherichia
129	25	80.6	633	1	CNE7_HUMAN	Q99d16	homo sapien	202	24	77.4	271	1	KKAI_SALTY	Q03447	salmoneilla
130	25	80.6	716	1	SYEC_SCHPO	O13775	schizosacch	203	24	77.4	272	1	KKAB_ECOLI	P14509	escherichia
131	25	80.6	727	1	NDAM_BOVIN	P15690	bos taurus	204	24	77.4	272	1	TFP1_HAEIN	P20588	haemophilus
132	25	80.6	883	1	RPOL_BPT7	P00573	bacterioph	205	24	77.4	280	1	TRPA_CORGL	P06562	corynebacte
133	25	80.6	884	1	PR06_YEAST	P19735	saccharomyc	206	24	77.4	288	1	LPXC_RICCN	O921t1	ricketsia
134	25	80.6	928	1	PR06_YEAST	P40048	saccharomyc	207	24	77.4	291	1	MSAB_TREPA	O83641	t.peptide m
135	25	80.6	928	1	PR06_YEAST	P40048	saccharomyc	208	24	77.4	293	1	YMO2_RHIME	P43906	rhizobium m
136	25	80.6	979	1	PL15_MYCHR	P41508	mycoplasma	209	24	77.4	296	1	ILVR_CAUCR	P52670	caulobacter
137	25	80.6	988	1	SY1_SYNY3	P73505	synchocyst	210	24	77.4	300	1	RIP3_MAIZE	P25891	zea mays (m
138	25	80.6	1013	1	YT93_DEIRA	Q97y19	deinococcu	211	24	77.4	301	1	RIPX_MAIZE	P25892	zea mays (m
139	25	80.6	1095	1	CC25_SACKL	O02342	saccharomyc	212	24	77.4	303	1	LPXC_PSEAE	P47205	pseudomonas
140	25	80.6	1249	1	TPP2_HUMAN	P29144	homo sapien	213	24	77.4	304	1	RIP9_MAIZE	P25892	zea mays (m
141	25	80.6	1249	1	TPP2_RAT	O64560	rattus norv	214	24	77.4	306	1	LPXC_SHEVI	O96v47	shevanelia
142	25	80.6	1262	1	TPP2_MOUSE	O64514	mus musculu	215	24	77.4	306	1	RDGC_PSEAE	O96v47	shevanelia
143	25	80.6	1302	1	UBPY_CAEEL	O09931	caenorhabdi	216	24	77.4	309	1	SANT_PLAFN	P04928	plasmidum
144	25	80.6	1331	1	XDH_BOVIN	R80457	bos taurus	217	24	77.4	321	1	VF05_YACCC	P21014	vaccinia vi
145	25	80.6	1537	1	DNM1_CHICK	Q92072	gallus gall	218	24	77.4	321	1	VF05_YACCP	Q00320	vaccinia vi
146	25	80.6	1690	1	RPOC_THEMA	P36252	thermotoga	219	24	77.4	322	1	VF05_YACCV	P24358	vaccinia vi
147	25	80.6	1748	1	YNR2_YEAST	P53886	saccharomyc	220	24	77.4	322	1	VP05_YARV	P33865	variola vir
148	25	80.6	1803	1	YU13_YEAST	P47024	saccharomyc	221	24	77.4	326	1	HEM2_PASMT	P57874	pasteurella
149	24	77.4	83	1	YB62_METUA	O58857	methanococc	222	24	77.4	333	1	HYU1_LACHE	P22294	lactobacilli
150	24	77.4	102	1	RPOL_AERPE	Q9ye21	aeropyrum p	223	24	77.4	342	1	YEA4_YEAST	P40004	saccharomyc
151	24	77.4	110	1	VCAD_LAMB	P03712	bacterioph	224	24	77.4	343	1	YNA1_ECOLI	P77253	escherichia
152	24	77.4	127	1	PAND_ANASP	O8y1r79	anabaena sp	225	24	77.4	355	1	BOK1_CLOAB	Q45829	clostridium
153	24	77.4	147	1	RUL3_STRCO	O53874	streptomyce	226	24	77.4	359	1	MSAB_HELPJ	Q9zmk8	h.peptide m
154	24	77.4	153	1	MLE_CAEEL	P53014	caenorhabdi	227	24	77.4	359	1	MSAB_HELPY	O25011	h.peptide m
155	24	77.4	156	1	MLE_AEOTR	P07291	aequipecten	228	24	77.4	360	1	VIS_AGRVT	O04554	agrobacteri
156	24	77.4	156	1	MLE_PATRE	P07280	palinopecten	229	24	77.4	374	1	HS83_DROVI	P04811	drosophila
157	24	77.4	171	1	NIDW_HUMAN	O96000	homo sapien	230	24	77.4	375	1	HS83_DROPS	P04809	drosophila
158	24	77.4	171	1	AHPD_MYCTU	O57353	mycobacteri	231	24	77.4	375	1	HS83_DROSI	P04810	drosophila
159	24	77.4	185	1	THF2_ARATH	O9xfn9	arabidopsi	232	24	77.4	375	1	SANT_PLAFV	P09593	plasmidum
160	24	77.4	188	1	Y084_NPVAC	P14744	autographa	233	24	77.4	390	1	MEIL_HUMAN	O00470	homo sapien
161	24	77.4	190	1	NUGM_ARATH	O95748	arabidopsi	234	24	77.4	390	1	MEIL_MOUSE	O60954	homo sapien
162	24	77.4	190	1	NUGM_ORSYA	Q35322	oryza sativ	235	24	77.4	390	1	MEIL_XENLA	P75937	xenopus lae
163	24	77.4	190	1	NUGM_SOLTU	R80261	solanum tub	236	24	77.4	391	1	MEIK_XENLA	O9s680	rhizobium l
164	24	77.4	192	1	NUGM_BETTR	Q33994	beta tritgyn	237	24	77.4	395	1	NDK7_MOUSE	O9qxl8	homo sapien
165	24	77.4	192	1	NUGM_BETVU	Q3787	beta vulgar	238	24	77.4	395	1	NDK7_RAT	O9qxl7	rattus norv
166	24	77.4	192	1	NUGM_BETWE	Q34011	beta webbia	239	24	77.4	398	1	DXR_ECO57	O8x8y1	escherichia
167	24	77.4	192	1	NUGM_PROMI	Q37622	protiotheca	240	24	77.4	398	1	DXR_ECOLI	P45568	escherichia
168	24	77.4	192	1	SCPA_PENSP	P02636	penaeus sp.	241	24	77.4	398	1	DXR_SALTI	O8x9a6	salmoneilla
169	24	77.4	193	1	YAJB_ECOLI	P21515	escherichia	242	24	77.4	398	1	DXR_SALTY	O8znp3	salmoneilla
170	24	77.4	195	1	NUGM_MARPO	P34944	marichantia	243	24	77.4	398	1	FGK_STRPN	O97889	streptococc
171	24	77.4	199	1	TPM1_YEAST	P17536	saccharomyc	244	24	77.4	399	1	FAST_RHOFA	P46373	rhodococcus
172	24	77.4	204	1	NUGM_RECAM	O21271	reclinomona	245	24	77.4	400	1	TRA6_BACST	Q45618	bacillus st
173	24	77.4	212	1	MSRA_SALTI	O82150	salmonella	246	24	77.4	405	1	CPYK_SACER	P33371	saccharopol
174	24	77.4	213	1	ID1_HAHLN1	O9h540	halobacteri	247	24	77.4	405	1	YLF2_YEAST	P338746	saccharomyc
175	24	77.4	215	1	FUCA_ECOLI	P19560	escherichia	248	24	77.4	409	1	V291_METUA	O57739	methanococc
176	24	77.4	217	1	PSBP_BRAJU	Q96334	brassica ju	249	24	77.4	410	1	SAHH_THEAC	O9hxx4	thermoplasm
177	24	77.4	228	1	YB62_CAEEL	Q21863	caenorhabdi	250	24	77.4	410	1	SAHH_THREVO	O97924	thermoplasm
178	24	77.4	232	1	YB62_CAEEL	Q21863	caenorhabdi	251	24	77.4	417	1	SYS_MYCGE	P47251	mycoplasma
179	24	77.4	234	1	INF8_HUMAN	P32971	homo sapien	252	24	77.4	424	1	VG46_BPPF1	P25131	bacterioph

253	24	77.4	425	1	PURA_FUSNN	P58793	fusobacteri	326	24	77.4	723	1	HS9B_RAT	P34058	rattus norv
254	24	77.4	425	1	PURA_FUSNB	O6581	fusobacteri	327	24	77.4	725	1	HS9B_BRARE	O57521	brachydanio
255	24	77.4	428	1	DHEA_SYNX3	P5386	synheocyst	328	24	77.4	725	1	HS9B_CHICK	O04619	gallus gall
256	24	77.4	432	1	PURA_YEAST	P80210	saccharomyc	329	24	77.4	726	1	HS9A_BRARE	O00474	brachydanio
257	24	77.4	443	1	MCRB_MEIVA	P07936	methanococc	330	24	77.4	728	1	HS9A_CHICK	P11501	gallus gall
258	24	77.4	443	1	MCRB_MEIVO	P11561	methanococc	331	24	77.4	732	1	ACET_HUMAN	P22966	homo sapien
259	24	77.4	443	1	PC17_HUMAN	Q01850	homo sapien	332	24	77.4	732	1	HS9A_CRIGR	P66633	cricetulus
260	24	77.4	451	1	CLUS_COTJA	P14018	coturnix co	333	24	77.4	732	1	HS9A_MOUSE	P07901	mus musculu
261	24	77.4	458	1	CHPN_HUMAN	P15169	homo sapien	334	24	77.4	732	1	HS9A_PIG	O02705	sus scrofa
262	24	77.4	458	1	MBPB_ENTHR	O08637	enterococcu	335	24	77.4	775	1	MCMS_YEAST	P29468	saccharomyc
263	24	77.4	464	1	RCCL_CANAL	P53439	candida alb	336	24	77.4	776	1	CG2E_MOUSE	P51944	mus musculu
264	24	77.4	465	1	HKK4_HUMAN	P35537	homo sapien	337	24	77.4	783	1	TRE2_THEBC	O33634	gallus gall
265	24	77.4	465	1	HKK4_MOUSE	P53792	mus musculu	338	24	77.4	785	1	CAD7_CHICK	Q07633	gallus gall
266	24	77.4	470	1	GLNA_METCA	P15124	methylococc	339	24	77.4	785	1	CAD7_HUMAN	Q01015	homo sapien
267	24	77.4	470	1	YMPB_CAEL	P53993	caenorhabdi	340	24	77.4	786	1	CG2F_HUMAN	P41003	homo sapien
268	24	77.4	471	1	MTHB_DROME	P83118	drosophila	341	24	77.4	789	1	K6PE_HABCO	O27665	haemophilu
269	24	77.4	472	1	ARPB_PERS	O50341	fervidobact	342	24	77.4	790	1	YDDB_ECOLI	P31827	escherichia
270	24	77.4	472	1	CBIR_FELCA	O02777	fells silve	343	24	77.4	801	1	CADK_HUMAN	Q48106	homo sapien
271	24	77.4	472	1	CBIR_HUMAN	P21534	homo sapien	344	24	77.4	862	1	ADHE_CLOAB	O33744	clostridium
272	24	77.4	472	1	ELBL_ADEA1	P10546	human adeno	345	24	77.4	901	1	GPOL_BORBU	O51157	borrelia bu
273	24	77.4	473	1	CBIR_MOUSE	P47746	mus musculu	346	24	77.4	908	1	ITP2_HUMAN	P19483	homo sapien
274	24	77.4	473	1	CBIR_POEGU	P56971	poephila gu	347	24	77.4	946	1	CSEL_YEAST	P33307	saccharomyc
275	24	77.4	473	1	CBIR_RAT	P20272	rattus norv	348	24	77.4	960	1	SUDJ_CAREL	P46549	caenorhabdi
276	24	77.4	497	1	Y124_YEAST	P47114	saccharomyc	349	24	77.4	982	1	TLR8_MOUSE	P38682	mus musculu
277	24	77.4	508	1	NOG0_DROME	O44081	drosophila	350	24	77.4	1032	1	RRPO_NODAV	Q31mm4	nodamura vi
278	24	77.4	508	1	YBCK_ECOLI	P77698	escherichia	351	24	77.4	1043	1	YBK1_YEAST	P3167	saccharomyc
279	24	77.4	509	1	DKC1_RAT	P4615	rattus norv	352	24	77.4	1077	1	SP23_YEAST	P52440	saccharomyc
280	24	77.4	514	1	DKC1_HUMAN	O60832	homo sapien	353	24	77.4	1082	1	RNT1_FUGRU	O9813	fugu rubrip
281	24	77.4	515	1	PEN3_ADEG1	O64755	avian adeno	354	24	77.4	1097	1	UBPB_SCHPO	O09879	schizosacch
282	24	77.4	517	1	UZAF_SCHPO	P36629	schizosacch	355	24	77.4	1108	1	MG42_YEAST	P40578	saccharomyc
283	24	77.4	521	1	CEAL_SHISO	P21178	shigella so	356	24	77.4	1113	1	RNT1_MOUSE	Q9epu	mus musculu
284	24	77.4	521	1	RE3_STRAU	O86490	staphylococ	357	24	77.4	1113	1	RNT1_HUMAN	Q92900	homo sapien
285	24	77.4	524	1	CP51_UNCNE	O14442	uncinula ne	358	24	77.4	1129	1	Y623_DROME	Q6767	drosophila
286	24	77.4	524	1	PPCK_CAMBE	O9pp01	campylobact	359	24	77.4	1145	1	ALAC_ARATH	O03792	arabidopsis
287	24	77.4	525	1	PEN3_ADEGX	P35358	avian adeno	360	24	77.4	1184	1	PROB_THEAC	O05855	mus musculu
288	24	77.4	531	1	UD16_RABIT	Q28611	oryctolagus	361	24	77.4	1195	1	FMN1_MOUSE	O05855	mus musculu
289	24	77.4	534	1	CN9A_MOUSE	O70628	mus musculu	362	24	77.4	1206	1	FMN1_CHICK	O05855	mus musculu
290	24	77.4	536	1	SYEM_YEAST	P48525	saccharomyc	363	24	77.4	1213	1	DBP2_KLULA	O42726	kluyveromyc
291	24	77.4	542	1	Y159_METTA	Q57623	methanococc	364	24	77.4	1221	1	REP2_CYAPA	Q9qz87	mus musculu
292	24	77.4	546	1	ALKK_PSEOL	Q00594	pseudomonas	365	24	77.4	1242	1	REOD_NODAV	P68120	cyranophora
293	24	77.4	548	1	CH60_ZYMO	P48220	zymomonas m	366	24	77.4	1265	1	BCSC_PSEPL	P58337	pseudomonas
294	24	77.4	550	1	CYRS_DICD1	P21837	dictyosteli	367	24	77.4	1279	1	ACE_HUMAN	P12821	homo sapien
295	24	77.4	558	1	RUBB_BRANA	P21241	brassica na	368	24	77.4	1306	1	VIVD_BPT7	P22986	bacterioph
296	24	77.4	589	1	IEP2_THEAC	Q9h160	thermoplasm	369	24	77.4	1318	1	KDH_RAT	O00519	mus musculu
297	24	77.4	589	1	IEP2_THEYO	Q97bK4	thermoplasm	370	24	77.4	1330	1	KDH_MOUSE	P37199	rattus norv
298	24	77.4	589	1	IEPB_ASCSU	P23731	ascaris suu	371	24	77.4	1335	1	N155_RAT	O75694	homo sapien
299	24	77.4	593	1	CN9A_HUMAN	O76083	homo sapien	372	24	77.4	1390	1	N155_HUMAN	P37199	rattus norv
300	24	77.4	593	1	SANT_PLAF7	O03400	plasmodium	373	24	77.4	1391	1	N155_MOUSE	O09988	mus musculu
301	24	77.4	595	1	RUBB_PEA	P06927	plum sativ	374	24	77.4	1391	1	SPT5_YEAST	P23615	saccharomyc
302	24	77.4	597	1	TYPA_SYNX3	P72749	synheocyst	375	24	77.4	1451	1	FMN1_MOUSE	O05860	mus musculu
303	24	77.4	600	1	RUBB_ARATH	P21240	arabidopsis	376	24	77.4	1468	1	YTO9_YEAST	P40453	saccharomyc
304	24	77.4	607	1	LPXB_CHLMD	O9hjy4	chlamydia m	377	24	77.4	1679	1	ZOI_HUMAN	O0715	homo sapien
305	24	77.4	613	1	DP2S_PTYRO	P81412	pyrococcus	378	24	77.4	1736	1	MYSA_RAT	Q04y43	rattus norv
306	24	77.4	619	1	PPCK_PYRAB	Q9uy53	pyrococcus	379	24	77.4	1828	1	MYSA_CHICK	O02440	gallus gall
307	24	77.4	618	1	SEUB_HAEIN	P43927	haemophilus	380	24	77.4	1829	1	MYSA_MOUSE	O09410	mus musculu
308	24	77.4	631	1	TAC3_MOUSE	Q9j111	mus musculu	381	24	77.4	1853	1	MYSA_HUMAN	O9y411	homo sapien
309	24	77.4	637	1	PRIM_LACLA	O04505	lactococcus	382	24	77.4	1855	1	T2D1_HUMAN	P21677	homo sapien
310	24	77.4	638	1	YOGS_BACSU	P54496	bacillus su	383	24	77.4	1872	1	POLG_CXA21	P22055	c genome po
311	24	77.4	664	1	UYRB_THEMA	O9y282	thermotoga	384	24	77.4	2206	1	RRPB_BEV	P22055	berne virus
312	24	77.4	668	1	DNJL_BACSU	O34398	bacillus su	385	24	77.4	2291	1	BPAL_HUMAN	O03001	homo sapien
313	24	77.4	670	1	CUI3_HUMAN	O9y447	homo sapien	386	24	77.4	2704	1	NOM1_YEAST	O00402	saccharomyc
314	24	77.4	685	1	RNT1_DROME	O9y453	drosophila	387	24	77.4	2748	1	FBNI_BOVIN	P58133	bos taurus
315	24	77.4	686	1	RECG_BORBU	O8y528	borrelia bu	388	24	77.4	2871	1	FBNI_HUMAN	P35555	homo sapien
316	24	77.4	693	1	RECG_ECOS7	O8y528	escherichia	389	24	77.4	2871	1	Q9tvt8	SUS	scrofa
317	24	77.4	693	1	RECG_ECOLI	P24230	escherichia	390	24	77.4	2871	1	CENF_HUMAN	P27395	j genome po
318	24	77.4	701	1	HS9D_PODAN	O43109	podospira a	391	24	77.4	3210	1	POLG_JAEV1	P19110	j genome po
319	24	77.4	709	1	CDAT_PLAFA	P46468	plasmodium	392	24	77.4	3432	1	POLG_JAEV5	P32886	j genome po
320	24	77.4	713	1	HS9B_HORSE	O094x8	equus cabal	393	24	77.4	3432	1	SRR2_BACSU	P04747	baclillus su
321	24	77.4	716	1	HS83_DROAV	O02192	drosophila	394	24	77.4	3587	1	FIXS_RHIME	O81399	rhizobium m
322	24	77.4	717	1	HS83_DROME	O02828	drosophila	395	24	77.4	3587	1	Y063_ARCTU	O30173	archaeoglob
323	24	77.4	719	1	HS9A_HORSE	O9gkx7	equus cabal	396	24	77.4	61	1	YAIK_ECOLI	P08366	escherichia
324	24	77.4	723	1	HS9B_HUMAN	P08238	homo sapien	397	24	77.4	63	1			
325	24	77.4	723	1	HS9B_MOUSE	P11499	mus musculu	398	24	77.4	63	1			

399	23	74.2	67	1	RF1_STRGC	P47850 streptococ	472	23	74.2	284	1	S4A1_MOUSE	Q9d597 mus musculu
400	23	74.2	85	1	VG73_BPM5	O65288 mycobacteri	473	23	74.2	284	1	TPM1_SCHMA	P42637 schistosoma
401	23	74.2	87	1	RI15_TOBAC	P6373 nicotiana t	474	23	74.2	290	1	ALP2_BACSU	P42420 bacillus su
402	23	74.2	88	1	RI15_ARATH	P68805 arabidopsis	475	23	74.2	295	1	OTC_HALNI	O48296 halobacteri
403	23	74.2	88	1	V176_TREPA	O83234 treponema p	476	23	74.2	305	1	CMRQ_SCHPO	O94417 schizosacch
404	23	74.2	90	1	RI15_SPTOL	O9m334 spiniacia ol	477	23	74.2	305	1	YES5_METUA	O58850 methanococ
405	23	74.2	96	1	NEUT_BRARE	Q31893 brachydanio	478	23	74.2	308	1	AMID_STEPN	P18794 streptococ
406	23	74.2	96	1	NEUT_YEAST	P28672 carassius a	479	23	74.2	309	1	WR26_ARATH	P51593 rattus norv
407	23	74.2	102	1	YNI6_YEAST	P28622 saccharomyc	480	23	74.2	310	1	UR21_RAT	P52836 flaveria ch
408	23	74.2	109	1	Y819_METUA	O58229 methanococ	481	23	74.2	311	1	F35F_FLACH	P52335 flaveria bi
409	23	74.2	120	1	RI17_BACHD	O50655 bacillus ha	482	23	74.2	312	1	F35F_FLACH	P52335 flaveria bi
410	23	74.2	121	1	RI14_SPTOL	P09596 spiniacia ol	483	23	74.2	312	1	YFCH_HAEIN	P45106 haemophilus
411	23	74.2	122	1	RI14_ARATH	P56792 arabidopsis ol	484	23	74.2	315	1	P2A_PARTE	P48726 parametium
412	23	74.2	123	1	RI14_MATZE	P08559 zea mays (m	485	23	74.2	317	1	OMPT_ECO57	P58603 escherichia
413	23	74.2	123	1	RI14_ORYSA	P12137 oryza sativ	486	23	74.2	321	1	OMPT_ECOLI	P09169 escherichia
414	23	74.2	123	1	RI14_TOBAC	P63342 nicotiana t	487	23	74.2	321	1	DE76_CHLPR	O9xly6 chlorella p
415	23	74.2	124	1	CD59_RABIT	O77541 cyclocolagus	488	23	74.2	322	1	XCX1_MOUSE	O9cm1 mus musculu
416	23	74.2	126	1	YD43_MYCLE	P54114 mycobacteri	489	23	74.2	322	1	YOUN_BACSU	P54338 bacillus su
417	23	74.2	126	1	YD43_MYCPT	O11013 mycobacteri	490	23	74.2	329	1	I329_ASEB7	P27345 african swi
418	23	74.2	136	1	RI17_RICPR	O9zc0 rickettsia	491	23	74.2	334	1	OOR_YEAST	P38230 saccharomyc
419	23	74.2	137	1	FLAY_MSEGL	P00321 megasphaera	492	23	74.2	336	1	RPOA_AGRIS	O9710 agrobacteri
420	23	74.2	141	1	RFH_ECOLI	P28359 escherichia	493	23	74.2	336	1	RPOA_RHILLO	O98h33 rhizobium l
421	23	74.2	157	1	YB23_PYRAB	Q9ux4 pyrococcus	494	23	74.2	337	1	RPOA_RHIME	O92522 rhizobium m
422	23	74.2	161	1	YY22_MYCLE	O49864 mycobacteri	495	23	74.2	337	1	RPOA_BRUME	O9d6c7 bartonella
423	23	74.2	161	1	IPYR_BACP3	P19514 bacillus ps	496	23	74.2	337	1	RPOA_BRUME	O8y16 bruceella me
424	23	74.2	164	1	IPYR_BACST	O05724 bacillus st	497	23	74.2	338	1	Y941_METUA	O57711 methanococ
425	23	74.2	166	1	VE07_VACCV	P21048 vaccinia vi	498	23	74.2	340	1	RPOA_RICCN	O92420 rickettsia
426	23	74.2	176	1	IPYR_GLUOX	O05545 glucinobact	499	23	74.2	340	1	RPOA_RICPR	O92659 rickettsia
427	23	74.2	176	1	RIW6_THEMA	Q9x194 thermotoga	500	23	74.2	341	1	AGC_METUA	O92659 rickettsia
428	23	74.2	185	1	PAGC_SALTY	P23988 salmonella	501	23	74.2	341	1	MAG_DERFA	O58496 dematophag
429	23	74.2	185	1	RPO6_MYXVL	P18630 myxoma viru	502	23	74.2	347	1	DHAS_LEGPN	O31219 leijonella
430	23	74.2	186	1	RPO6_FOWPV	O85280 fowlpox vir	503	23	74.2	347	1	H806_METUA	O58116 leijonella
431	23	74.2	188	1	Y4A0_RHISM	P55364 rhizobium s	504	23	74.2	348	1	ARO_HRWEE	O54459 erwinia her
432	23	74.2	195	1	NUGM_ACACA	O37383 acanthamoeb	505	23	74.2	348	1	VMAT_SENDS	P27567 sendai viru
433	23	74.2	198	1	PBD2_CHICK	O73888 gallus gall	506	23	74.2	348	1	VMAT_SENDS	P17748 sendai viru
434	23	74.2	198	1	PBD2_CHICK	Q9jht7 mus musculu	507	23	74.2	348	1	VMAT_SENDS	P03426 sendai viru
435	23	74.2	198	1	PBD2_MOUSE	O55343 rattus norv	508	23	74.2	348	1	VMAT_SENDS	P06446 sendai viru
436	23	74.2	201	1	MUSC_MOUSE	O88940 mus musculu	509	23	74.2	349	1	AOX1_SAVGU	P22185 saurotatum
437	23	74.2	204	1	VATD_BORBU	O51119 borrelia bu	510	23	74.2	351	1	CBF1_YEAST	P17106 saccharomyc
438	23	74.2	207	1	RID2_BUCAL	P57534 buchera ap	511	23	74.2	351	1	VORH_METTH	O26800 methanobact
439	23	74.2	213	1	YIN2_YEAST	P40465 saccharomyc	512	23	74.2	353	1	HME1_DROME	P09089 dirosophila
440	23	74.2	215	1	TRJ9_ECOLI	Q00738 escherichia	513	23	74.2	353	1	IL6A_RABIT	P21109 oryctolagus
441	23	74.2	216	1	FUCA_HAEIN	P44777 haemophilus	514	23	74.2	357	1	RF1_LACLA	O9cmx3 lactococcus
442	23	74.2	218	1	RL10_DROME	O61231 dirosophila	515	23	74.2	358	1	PHLC_TREB	O9x194 trypanosoma
443	23	74.2	221	1	SFR9_HUMAN	O13242 homo sapien	516	23	74.2	359	1	RF1_STRAG	O9a229 streptococ
444	23	74.2	224	1	KCY_THETN	O8ra78 thermoaer	517	23	74.2	359	1	RF1_STRPN	O97420 streptococ
445	23	74.2	224	1	E681_CHLPH	O927m3 chlamydia p	518	23	74.2	359	1	RF1_STRPV	O9a925 streptococ
446	23	74.2	232	1	E4RE_METHN	O26350 methanobact	519	23	74.2	360	1	IL6B_BOVIN	O28603 bos taurus
447	23	74.2	232	1	YB9H_YEAST	P38340 saccharomyc	520	23	74.2	360	1	RF1_PSEAE	P42806 pseudomonas
448	23	74.2	237	1	VD03_VACCC	P21009 vaccinia vi	521	23	74.2	361	1	RF1_HAEIN	O90336 cyprinus ca
449	23	74.2	237	1	VD03_VACCV	P43032 vaccinia vi	522	23	74.2	361	1	RF1_HAEIN	P43917 haemophilus
450	23	74.2	237	1	YFBN_ECOLI	P33068 variola vir	523	23	74.2	366	1	H182_HAEIN	O57004 haemophilus
451	23	74.2	238	1	YFBN_ECOLI	P6484 escherichia	524	23	74.2	368	1	DXR_HELPJ	O9m16 helicobacte
452	23	74.2	243	1	GNTB_BACSU	P10585 bacillus su	525	23	74.2	370	1	SNCL_HUMAN	O16533 homo sapien
453	23	74.2	244	1	WC23_VACCC	P21090 vaccinia vi	526	23	74.2	370	1	K2C8_BOVIN	P05786 bos taurus
454	23	74.2	250	1	TDXH_AERPE	O9y910 aeropyrum p	527	23	74.2	372	1	CARA_CAMDE	O9pmg8 campylobact
455	23	74.2	253	1	HIS6_PYRAE	O8zy16 pyrobaculum	528	23	74.2	373	1	HME1_DROSU	O24648 dirosophila
456	23	74.2	253	1	VC23_VAVV	P34016 variola vir	529	23	74.2	374	1	YG09_METUA	O59004 methanococ
457	23	74.2	257	1	ICEE_MOUSE	O89094 mus musculu	530	23	74.2	375	1	YHR4_PSEY	P20400 pseudomonas
458	23	74.2	258	1	VC23_VACCC	P10663 vaccinia vi	531	23	74.2	384	1	O46B_DROME	O9y3n2 dirosophila
459	23	74.2	261	1	TNE5_BOVIN	P51749 bos taurus	532	23	74.2	389	1	AAT_SYNY3	O55128 synechocyst
460	23	74.2	262	1	TRPA_AQUAE	O67502 aquifex aeo	533	23	74.2	393	1	PGIR_PYRPE	P48979 prunus pers
461	23	74.2	264	1	TYST_RALISO	O8y06 ralsomnia s	534	23	74.2	394	1	EFPU_BORBU	P50062 borrelia bu
462	23	74.2	265	1	HME2_BRARE	P09015 brachydanio	535	23	74.2	398	1	PSMR_ARCFU	O28303 archaeglob
463	23	74.2	270	1	YN8A_YEAST	P53722 saccharomyc	536	23	74.2	398	1	Y4QJ_RHISM	P55631 rhizobium s
464	23	74.2	271	1	PUNA_BACSU	P46354 bacillus su	537	23	74.2	400	1	YD01_METUA	O58697 methanococ
465	23	74.2	272	1	VG15_BP4	P11112 bacterioph	538	23	74.2	404	1	REPA_AGRRH	P05682 agrobacteri
466	23	74.2	276	1	PRXC_SYNY3	O55921 synechocyst	539	23	74.2	406	1	OLIA_DROME	O9y221 dirosophila
467	23	74.2	278	1	UL07_EBY	P29882 epstein-bar	540	23	74.2	410	1	FBW4_MOUSE	O9ymj2 mus musculu
468	23	74.2	279	1	YEB8_YEEST	P39995 saccharomyc	541	23	74.2	411	1	INTR_ECOLI	P76056 escherichia
469	23	74.2	282	1	TR2A_HUMAN	O13595 homo sapien	542	23	74.2	412	1	FBW4_HUMAN	P57775 homo sapien
470	23	74.2	283	1	RUPAP_SOYBN	P39657 glycine max	543	23	74.2	413	1	DCHS_LYCES	P54772 lycopersico
471	23	74.2	284	1	S4A1_HUMAN	O9br01 homo sapien	544	23	74.2	413	1	DP41_BACHD	O9kc07 bacillus ha

545	23	74.2	417	1	GLYA_YERPE	Q8crl1	yersinia pe	618	23	74.2	525	1	RRPO_IBDVS	P318:7	avian infec
546	23	74.2	419	1	NOEE_RHISN	P55472	rhizobium s	619	23	74.2	537	1	YCUB_SCHPO	O59831	schizosacch
547	23	74.2	420	1	Y103_SYNM3	O55874	synecocyst	620	23	74.2	541	1	HXT2_YEAST	P23595	saccharomyc
548	23	74.2	421	1	DCDA_ZYMO	O95661	zymomonas m	621	23	74.2	542	1	V155_FOMPV	O91587	fowlpox vit
549	23	74.2	422	1	RPC4_YEAST	P25441	saccharomyc	622	23	74.2	545	1	MUTL_THETH	O9ra54	thermus the
550	23	74.2	426	1	SYS_TREPA	O83653	treponema p	623	23	74.2	546	1	SYR_THEMA	Q9X018	thermotoga
551	23	74.2	432	1	DADA_ECOLI	P23011	escherichia	624	23	74.2	546	1	TCPE_SCHPO	Q9tun4	schizosacch
552	23	74.2	432	1	HSD_PSEAE	O9x490	pseudomonas	625	23	74.2	554	1	CH60_HOLOB	P44820	holospira o
553	23	74.2	432	1	OAT_DROAN	P49724	drosophila	626	23	74.2	556	1	YDPA_SCHPO	O14013	schizosacch
554	23	74.2	435	1	ORC6_YEAST	P38826	saccharomyc	627	23	74.2	565	1	PRIM_THEMA	O7x123	thermotoga
555	23	74.2	436	1	CCKR_MOUSE	O08786	mus musculus	628	23	74.2	566	1	HS60_CANAL	O74261	candida alb
556	23	74.2	437	1	AP2A_HUMAN	P05549	homo sapien	629	23	74.2	578	1	TRM1_DROME	O9vkr9	drosophila
557	23	74.2	437	1	AP2A_MOUSE	P34056	mus musculus	630	23	74.2	586	1	SYR_STRCO	Q9x423	streptomyce
558	23	74.2	437	1	AP2A_RAT	P38197	rattus norv	631	23	74.2	587	1	G6T_BACSU	P34422	bacillus su
559	23	74.2	440	1	GP02_YEAST	P41911	saccharomyc	632	23	74.2	588	1	VCLB_GOSHI	P09801	gossypium h
560	23	74.2	444	1	CKR1_RAT	P30551	rattus norv	633	23	74.2	591	1	YN48_YEAST	P24846	saccharomyc
561	23	74.2	450	1	SWP1_ENCCU	O9x2v1	encephalito	634	23	74.2	597	1	BP10_PARLI	P42670	paracentrot
562	23	74.2	453	1	UCR2_HUMAN	O94b77	homo sapien	635	23	74.2	597	1	SYK_AQUAE	O67258	aquifex aeo
563	23	74.2	453	1	UCR2_MOUSE	O94b77	mus musculus	636	23	74.2	602	1	LEPA_HELPI	O25122	heliobacte
564	23	74.2	454	1	CBPH_LOPAM	P37892	lophius ame	637	23	74.2	604	1	LEPA_HELPI	O9zmr3	heliobacte
565	23	74.2	458	1	ARLY_BACD	O94821	bacillus ha	638	23	74.2	605	1	VCLA_GOSHI	P09759	gossypium h
566	23	74.2	462	1	ATPB_CHLI	P35110	chlorobium	639	23	74.2	609	1	ALBU_MERUN	O35050	meriones un
567	23	74.2	462	1	ATPB_CHLI	P42465	chlorobium	640	23	74.2	611	1	IF4B_HUMAN	P23588	homo sapien
568	23	74.2	464	1	YCF1_MARPO	P12223	marcattia	641	23	74.2	612	1	HS76_YEAST	P11484	saccharomyc
569	23	74.2	465	1	REPA_VIBCH	O07024	vibrio chol	642	23	74.2	612	1	HS76_YEAST	P40150	saccharomyc
570	23	74.2	467	1	ISP6_SCHPO	P40903	schizosacch	643	23	74.2	613	1	ENP4_MOUSE	O9btr4	mus musculu
571	23	74.2	467	1	LIPI_HUMAN	P54315	homo sapien	644	23	74.2	613	1	HS75_SCHPO	O10284	schizosacch
572	23	74.2	467	1	XAMB_XANCP	P29356	x xanthan b	645	23	74.2	616	1	ENP4_HUMAN	O9y227	homo sapien
573	23	74.2	468	1	ATPB_THEMA	O50550	thermotoga	646	23	74.2	620	1	Y512_CHLPR	Q92840	chlamydia p
574	23	74.2	469	1	GINA_PROVU	P28786	proteus vul	647	23	74.2	621	1	Y425_CHLPR	O84432	chlamydia t
575	23	74.2	470	1	VATB_ARCFU	O29100	archaeoglob	648	23	74.2	621	1	Y708_CHLMU	O93j48	chlamydia s
576	23	74.2	470	1	YATB_YEAST	O04371	saccharomyc	649	23	74.2	622	1	PPCK_RALSO	O8v338	ralstonia s
577	23	74.2	471	1	APPB_HERAU	P47466	herpetosiph	650	23	74.2	628	1	MXFI_CABEL	O43767	caenorhabdi
578	23	74.2	473	1	APPB_BACSU	P37809	bacillus su	651	23	74.2	629	1	RRSD_HAEIN	P43766	caenorhabdi
579	23	74.2	473	1	CB1R_TARGR	O9pui7	taticia gra	652	23	74.2	633	1	DNAR_AGRIS	P30019	haemophilus
580	23	74.2	473	1	EX1_HAEIN	P45188	haemophilus	653	23	74.2	633	1	RRSD_RICPE	P33451	ricketsia
581	23	74.2	474	1	PRTP_MOUSE	P16675	mus musculu	654	23	74.2	635	1	DNAR_RHILE	O33528	rhizobium 1
582	23	74.2	475	1	VATB_MANSE	O91550	manduca sex	655	23	74.2	638	1	HS74_PARLI	O62428	paracentrot
583	23	74.2	477	1	MYPH_HUMAN	O13203	homo sapien	656	23	74.2	643	1	YNV4_CABEL	P45657	caenorhabdi
584	23	74.2	480	1	MANC_SALTY	P26340	salmonella	657	23	74.2	644	1	VHJ9_YEAST	P38664	saccharomyc
585	23	74.2	480	1	SES2_HUMAN	P56004	homo sapien	658	23	74.2	645	1	LIPI_PHOTO	P40650	photorhabdu
586	23	74.2	480	1	SES2_MOUSE	P56043	mus musculu	659	23	74.2	647	1	VEI_HPV52	P36730	human papil
587	23	74.2	481	1	PHR_HALYA	P20077	halobacteri	660	23	74.2	647	1	PIGI_YEAST	O66232	saccharomyc
588	23	74.2	481	1	SESI_XENIA	P58033	xenopus lae	661	23	74.2	658	1	HS7F_CABEL	P11141	caenorhabdi
589	23	74.2	482	1	G3B2_HUMAN	O9un86	homo sapien	662	23	74.2	659	1	GUN3_FIBSU	P42650	fibrobacteri
590	23	74.2	482	1	G3B2_MOUSE	P97379	mus musculu	663	23	74.2	659	1	MUTL_HALNI	O9h560	halobacteri
591	23	74.2	482	1	K2C8_HUMAN	P05787	homo sapien	664	23	74.2	671	1	RECG_STRPN	O54900	streptococc
592	23	74.2	482	1	K2C8_RAT	O10758	rattus norv	665	23	74.2	681	1	CRY1_ARATH	O43125	strabidopsis
593	23	74.2	485	1	ATPB_MYCLE	P45823	mycobacteri	666	23	74.2	686	1	RECG_STRAU	O30361	staphylococ
594	23	74.2	486	1	ATPB_MYCU	O10593	mycobacteri	667	23	74.2	692	1	PHLA_ECOLI	O80983	escherichia
595	23	74.2	486	1	K1C0_XENIA	P02537	xenopus lae	668	23	74.2	694	1	FRFS_YEAST	O08908	saccharomyc
596	23	74.2	488	1	K2C8_MOUSE	P11679	mus musculu	669	23	74.2	696	1	RIR2_AOUAE	O67475	aquifex aeo
597	23	74.2	489	1	PGI1_YEAST	P38720	saccharomyc	670	23	74.2	696	1	Y195_SYNV3	P41010	synecocyst
598	23	74.2	492	1	PGC2_YEAST	P53319	saccharomyc	671	23	74.2	698	1	Z234_HUMAN	O14588	homo sapien
599	23	74.2	492	1	SESI_HUMAN	O9y6p5	homo sapien	672	23	74.2	700	1	RIR1_BACSU	P08620	bacillus su
600	23	74.2	492	1	SESI_MOUSE	P56006	mus musculu	673	23	74.2	705	1	FB11_MOUSE	Q98879	mus musculu
601	23	74.2	492	1	SEB3_HUMAN	P56006	mus musculu	674	23	74.2	713	1	FP2_MOUSE	P20239	mus musculu
602	23	74.2	492	1	SEB3_MOUSE	O9cyp7	mus musculu	675	23	74.2	719	1	MOS2_STRAU	Q90259	staphylococ
603	23	74.2	493	1	TEBA_STYMY	P23550	stylyonchia	676	23	74.2	731	1	HS9A_HUMAN	O67960	homo sapien
604	23	74.2	495	1	NUSA_ECOLI	P03003	escherichia	677	23	74.2	732	1	ADDI_CABEL	O90900	caenorhabdi
605	23	74.2	495	1	TEBA_OXYNO	P23544	oxytricha n	678	23	74.2	748	1	PURL_DEIRA	O9xrt4	deinococcus
606	23	74.2	496	1	TEBA_THEMA	O9x0n9	thermotoga	679	23	74.2	748	1	CIPE_LACTA	O9c109	lactococcus
607	23	74.2	499	1	RUBB_SECC	O43831	secale cere	680	23	74.2	748	1	CIPE_LACTC	Q9s552	lactococcus
608	23	74.2	501	1	ATPA_ACEMO	P50000	acetobacter	681	23	74.2	748	1	STC_HUMAN	P49589	homo sapien
609	23	74.2	503	1	CUS_BACFI	O66043	bacillus fi	682	23	74.2	762	1	PMT4_YEAST	P46971	saccharomyc
610	23	74.2	504	1	Y4YA_RHISN	O55501	rhizobium s	683	23	74.2	768	1	DNL1_SCHPO	P12000	schizosacch
611	23	74.2	511	1	C716_NEPRA	O04164	nepeira race	684	23	74.2	775	1	TNP3_MOUSE	O60769	mus musculu
612	23	74.2	511	1	MDOG_ECOLI	P33136	escherichia	685	23	74.2	778	1	SEAD_MYCLE	O32922	mycobacteri
613	23	74.2	513	1	TPS1_SCHPO	P40387	schizosacch	686	23	74.2	780	1	K6PL_HUMAN	P12382	homo sapien
614	23	74.2	517	1	6RCD_CANAL	O13287	candida alb	687	23	74.2	780	1	K6PL_MOUSE	P12382	mus musculu
615	23	74.2	518	1	PTMB_STACA	P28008	staphylococ	688	23	74.2	780	1	K6PL_RAT	P30833	rattus norv
616	23	74.2	522	1	TYRI_CARAU	P55028	carassius a	689	23	74.2	782	1	MOS2_STAM	Q99uwl	staphylococ
617	23	74.2	524	1	YBY8_YEAST	P38277	saccharomyc	690	23	74.2	784	1	ALP4_SCHPO	O9y705	schizosacch

691	23	74.2	785	1	MUS2_LISTN	092c6	1	764	23	74.2	1756	1	PEPL_HUMAN	060437	homo sapien
692	23	74.2	785	1	MUS2_LISTO	0872p1	1	765	23	74.2	1790	1	SEPA_EMENT	P78621	emericeila
693	23	74.2	785	1	PKC7_HUMAN	Q16549	homo sapien	766	23	74.2	1822	1	ZAP3_HUMAN	P49750	homo sapien
694	23	74.2	790	1	TNP3_HUMAN	P21550	homo sapien	767	23	74.2	1969	1	MYSA_CAEEL	P12844	caenorhabdi
695	23	74.2	795	1	PEN_HUMAN	Q9ug90	homo sapien	768	23	74.2	2064	1	SIF1_DROME	P91621	drosophila
696	23	74.2	797	1	Y900_TREPA	083870	treponema p	769	23	74.2	2131	1	CINA_DROME	P33500	drosophila
697	23	74.2	799	1	CAD8_HUMAN	P52826	homo sapien	770	23	74.2	2163	1	BRR2_YEAST	P32635	saccharomyc
698	23	74.2	799	1	CAD8_MOUSE	P97291	mus musculus	771	23	74.2	2295	1	MDR3_HUMAN	Q9ns16	homo sapien
699	23	74.2	799	1	CAD8_RAT	054800	rattus norv	772	23	74.2	2366	1	TOXN_CLODI	P18177	clostridium
700	23	74.2	808	1	SEAD_MOUSE	050612	mycobacteri	773	23	74.2	2493	1	YBA4_YEAST	P35194	saccharomyc
701	23	74.2	810	1	V2A_BBYV	P27452	broad bean	774	23	74.2	2564	1	SPCO_HUMAN	Q9n254	homo sapien
702	23	74.2	824	1	SILP_SALTY	Q9zhc7	salmonella	775	23	74.2	3430	1	POLG_MNV	P06935	w genome po
703	23	74.2	837	1	YL34_YEAST	Q07844	saccharomyc	776	23	74.2	3433	1	POLG_KUNJM	P14335	k genome po
704	23	74.2	839	1	V2A_CMVO	P06012	cucumber m	777	23	74.2	4289	1	TENX_HUMAN	P22105	homo sapien
705	23	74.2	843	1	RPOC_OPOSI	P49487	odontella s	778	23	74.2	4344	1	DYHC_EMENT	P45454	emeriella
706	23	74.2	844	1	BPBA_ECOLI	P02919	e penicilli	779	23	74.2	4563	1	APR_HUMAN	P04114	homo sapien
707	23	74.2	855	1	BPBA_AQUAE	067889	aquifex aeo	780	23	74.2	4568	1	DYHC_CAEEL	P04114	homo sapien
708	23	74.2	868	1	MCM2_YEAST	P29469	saccharomyc	781	23	74.2	6609	1	NEBO_HUMAN	P20920	homo sapien
709	23	74.2	875	1	Z1P1_YEAST	P31111	saccharomyc	782	23	71.0	77	1	CPV9_BE1B6	P21682	bacterioph
710	23	74.2	878	1	RPO5_HUMAN	P12918	avian infec	783	23	71.0	99	1	YF71_HAEIN	P44260	streptomyc
711	23	74.2	879	1	ORP5_HUMAN	Q9h0x9	homo sapien	784	23	71.0	83	1	ACPX_SPRGA	P12884	streptomyc
712	23	74.2	903	1	VGILB_HSV1F	P06436	herpes simp	785	23	71.0	86	1	ACPX_SPRCO	002054	streptomyc
713	23	74.2	903	1	VB56_METUA	058556	methanococ	786	23	71.0	99	1	S214_HUMAN	095715	homo sapien
714	23	74.2	904	1	VGILB_HSV11	P10211	herpes simp	787	23	71.0	99	1	S214_MOUSE	09wq45	mus musculi
715	23	74.2	904	1	VGILB_HSV1K	P06437	herpes simp	788	23	71.0	101	1	IF52_CHICK	009121	gallus gall
716	23	74.2	904	1	VGILB_HSV1P	P06763	herpes simp	789	23	71.0	103	1	ACP_CYAPA	P48078	cyanophora
717	23	74.2	904	1	VGILB_HSV23	P08666	herpes simp	790	23	71.0	103	1	Y041_AQUAE	066457	aquiflex aeo
718	23	74.2	904	1	VGILB_HSV2H	Q9hml1	medicago tr	791	23	71.0	111	1	RPOL_TREAC	09h1c5	thermoplasm
719	23	74.2	935	1	IF38_MEDTR	P38250	saccharomyc	792	23	71.0	111	1	RPOL_TREAC	0978f5	thermoplasm
720	23	74.2	945	1	YBT6_YEAST	P07118	escherichia	793	23	71.0	111	1	Y065_BACAN	09m99	baacillus an
721	23	74.2	951	1	SVV_ECOLI	Q8kp73	vibrio chol	794	23	71.0	115	1	Y420_AQUAE	066736	aquiflex aeo
722	23	74.2	953	1	SVV_VITBCH	Q09457	caenorhabdi	795	23	71.0	115	1	YXBA_BACSU	P54803	baacillus su
723	23	74.2	963	1	YQ36_CAEEL	P02659	butyrylvi	796	23	71.0	116	1	RS24_METKA	0812d8	methanopyru
724	23	74.2	963	1	AMY_BUTFI	P03204	epstein-bar	797	23	71.0	120	1	HEMX_THETS	09xv29	thermomyzon
725	23	74.2	992	1	EBN6_EBV	Q96ma2	homo sapien	798	23	71.0	122	1	AC52_ECO57	08xan7	escherichia
726	23	74.2	994	1	NALA_HUMAN	Q20168	caenorhabdi	799	23	71.0	123	1	SVUG_MOUSE	09z0f7	mus musculi
727	23	74.2	1000	1	COPE_CAEEL	P20043	lactobacilli	800	23	71.0	123	1	SVUG_RAT	065544	rattus norv
728	23	74.2	1006	1	BGAL_LACDE	Q58863	methanococ	801	23	71.0	124	1	HYPA_METUA	057667	methanococ
729	23	74.2	1009	1	YB68_METUA	Q59140	arthrobaacte	802	23	71.0	127	1	STUG_HUMAN	076070	homo sapien
730	23	74.2	1013	1	BGAL_AR1SP	Q07722	bos taurus	803	23	71.0	130	1	CHEY_CAMJE	P71129	campylobact
731	23	74.2	1023	1	PIB4_BOVIN	Q9bxc6	homo sapien	804	23	71.0	130	1	RP5M_ALCJE	P28613	alcaligenes
732	23	74.2	1037	1	CAR6_HUMAN	Q104x6	schistosach	805	23	71.0	134	1	PRL4_HUMAN	Q16379	homo sapien
733	23	74.2	1040	1	RIK1_SCHPO	P16066	homo sapien	806	23	71.0	136	1	CT2A_MOUSE	P13399	mus musculi
734	23	74.2	1061	1	ANPA_HUMAN	P15938	saccharomyc	807	23	71.0	138	1	GTH2_CLAGA	P55453	clarifas gar
735	23	74.2	1071	1	PR16_YEAST	Q19192	caenorhabdi	808	23	71.0	140	1	Y850_AQUAE	067017	aquiflex aeo
736	23	74.2	1077	1	E2K3_CAEEL	Q99014	trichoderma	809	23	71.0	140	1	YAAH_SCHPO	Q08002	schistosach
737	23	74.2	1139	1	KPC1_TRIPE	P87253	neurospora	810	23	71.0	141	1	YOHJ_HAEIN	P45145	haemophilus
738	23	74.2	1142	1	KPC1_NEOCR	P24384	saccharomyc	811	23	71.0	141	1	CT2B_MOUSE	P14400	mus musculi
739	23	74.2	1145	1	PR22_YEAST	P14737	xanthomonas	812	23	71.0	142	1	GTH2_CORAU	P48251	coregonus a
740	23	74.2	1164	1	YAV1_XANCV	059037	methanococ	813	23	71.0	142	1	GTH2_ONCKE	P10256	oncorhynch
741	23	74.2	1169	1	SMC_METUA	Q15147	homo sapien	814	23	71.0	142	1	PSAD_MASTA	007115	maestriciad
742	23	74.2	1175	1	PIB4_HUMAN	Q9q407	rattus norv	815	23	71.0	144	1	AROO_CIOAB	Q97118	clostridium
743	23	74.2	1175	1	FIB4_RAT	Q9p943	xytella fas	816	23	71.0	144	1	RPOE_MYCPN	P75090	mycoplasma
744	23	74.2	1193	1	DP3A_XYLEFA	Q9p943	xytella fas	817	23	71.0	144	1	RPOE_MYCPN	P75090	mycoplasma
745	23	74.2	1213	1	T2D2_DROME	P42175	drosophila	818	23	71.0	145	1	FLAV_DESGI	P47268	mycoplasma
746	23	74.2	1228	1	NARG_BACSU	P42175	drosophila	819	23	71.0	146	1	YV46_HELPJ	Q01095	desulfovibr
747	23	74.2	1241	1	RPOB_CIOAB	P97499	clostridium	820	23	71.0	146	1	YV46_HELPJ	Q02687	heliobacte
748	23	74.2	1242	1	DPOL_HCMVA	P08546	human cytom	821	23	71.0	147	1	YBSE_BACST	031511	baacillus su
749	23	74.2	1266	1	YRD6_CAEEL	P09575	caenorhabdi	822	23	71.0	148	1	COX4_DICDI	P30815	dicyosteli
750	23	74.2	1274	1	BXF_CLOBO	P30936	clostridium	823	23	71.0	149	1	NDK1_ARATH	P39207	arabidopsi
751	23	74.2	1332	1	IKAP_HUMAN	Q95163	homo sapien	824	23	71.0	149	1	PRO5_ORYZA	P18085	oryza sativ
752	23	74.2	1335	1	RKPO_FKXAV	P22168	foxtail mos	825	23	71.0	151	1	REG6_PYRAB	Q9V0Y9	pyrococcus
753	23	74.2	1336	1	W146_HUMAN	Q9C018	homo sapien	826	23	71.0	152	1	YF63_METUA	Q58958	methanococ
754	23	74.2	1353	1	XDH_CALVI	P08793	calliphora	827	23	71.0	153	1	IF51_CHICK	Q07460	gallus gall
755	23	74.2	1356	1	XDH_CHICK	P47990	gallus gall	828	23	71.0	153	1	IF5A_HUMAN	P10159	homo sapien
756	23	74.2	1407	1	T2D2_YEAST	P23255	saccharomyc	829	23	71.0	153	1	IF5A_RABIT	P10160	oryctolagus
757	23	74.2	1501	1	EMBS_CAEEL	P34703	caenorhabdi	830	23	71.0	154	1	NDK_NEUCR	Q9Uuy8	neurospora
758	23	74.2	1572	1	BAI2_HUMAN	Q60241	homo sapien	831	23	71.0	157	1	IF51_SCHPO	P56289	schistosach
759	23	74.2	1585	1	YOBO_BACSU	P45941	baacillus su	832	23	71.0	157	1	IF52_SCHPO	Q9us14	schistosach
760	23	74.2	1654	1	PCFB_HUMAN	Q94913	homo sapien	833	23	71.0	160	1	ALTA_ECOLI	P77731	escherichia
761	23	74.2	1654	1	INT1_CANAL	P53705	candida alb	834	23	71.0	160	1	ALTA_ECOLI	Q944y6	salmonella
762	23	74.2	1750	1	Y832_METUA	058242	methanococ	835	23	71.0	161	1	NDK_HALNT	Q9ngns	halobacteri
763	23	74.2	1755	1	PEPL_MOUSE	Q9r289	mus musculi	836	23	71.0	162	1	IPYR_MYCLE	069540	mycobacteri

837	22	71.0	162	1	IPYR_MYCTU	006379 mycobacteri	910	22	71.0	227	1	COX2_BUBDE	P50678 bubalus dep
838	22	71.0	167	1	NDOE_RICCN	Q931d9 rickettsia	911	22	71.0	227	1	COX2_CANAD	O47667 canis adust
839	22	71.0	168	1	GA15_CRILLO	P14607 cricetus	912	22	71.0	227	1	COX2_CANAD	O47667 canis aureu
840	22	71.0	168	1	GA15_MOUSE	P35639 mus musculu	913	22	71.0	227	1	COX2_CANFA	O63855 canis famli
841	22	71.0	169	1	GA15_HUMAN	P35638 homo sapien	914	22	71.0	227	1	COX2_CANME	O47671 canis mesom
842	22	71.0	173	1	CD3D_RAT	P19377 rattus norv	915	22	71.0	227	1	COX2_CANST	P80831 canis simen
843	22	71.0	174	1	CO22_HOMMO	O02388 bombyx mori	916	22	71.0	227	1	COX2_CAPHI	O37430 capra hircu
844	22	71.0	174	1	GREP_METTH	O02350 methanobact	917	22	71.0	227	1	COX2_CERST	O03851 ceratotheri
845	22	71.0	175	1	CYT2_STOHE	P07845 stoichiactis	918	22	71.0	227	1	COX2_CERNR	O37448 cervus unic
846	22	71.0	175	1	NIDM_BOVIN	O02373 bos taurus	919	22	71.0	227	1	COX2_CHRRR	O47670 chrysoeyon
847	22	71.0	176	1	MOC_HELPU	O92145 helicobacte	920	22	71.0	227	1	COX2_CRAAU	P50685 cratogeomys
848	22	71.0	176	1	MOC_HELPU	P56421 helicobacte	921	22	71.0	227	1	COX2_CRAUA	O47668 cuon alpinu
849	22	71.0	177	1	FABA_PASMU	Q9cm62 pasteurella	922	22	71.0	227	1	COX2_CUCAL	P50679 damaliscus
850	22	71.0	177	1	NDOE_RICPR	Q9cm65 rickettsia	923	22	71.0	227	1	COX2_DAMP	P50687 dasypus nov
851	22	71.0	177	1	VC08_YACCV	P17364 vaccinia vi	924	22	71.0	227	1	COX2_DASNO	O47672 dulsioyon th
852	22	71.0	178	1	YEDR_ECOLI	P76514 escherichia	925	22	71.0	227	1	COX2_DUSNH	O47673 equus asinu
853	22	71.0	179	1	UBCB_XENLA	P56616 xenopus lae	926	22	71.0	227	1	COX2_EOUAS	P18890 felis silve
854	22	71.0	179	1	UBCC_HUMAN	O00762 homo sapien	927	22	71.0	227	1	COX2_FELCA	P48890 gazella spe
855	22	71.0	181	1	TCIP_BRUMA	P09697 brugia mala	928	22	71.0	227	1	COX2_GALSP	P38596 halichoerus
856	22	71.0	181	1	TCIP_MUCBA	Q962a2 wuchereria	929	22	71.0	227	1	COX2_HALGR	Q92zy9 hippopotamu
857	22	71.0	181	1	VHL_MOUSE	P40338 mus musculu	930	22	71.0	227	1	COX2_HIPAM	P48660 equus cabal
858	22	71.0	184	1	VC08_YACCC	P21041 vaccinia vi	931	22	71.0	227	1	COX2_HORSE	P98035 lemur caria
859	22	71.0	184	1	YARO_ACTPL	P22624 methanother	932	22	71.0	227	1	COX2_LEMCA	O47674 lycan pict
860	22	71.0	188	1	AMPN_METFE	P34264 caenorhabdi	933	22	71.0	227	1	COX2_LYCPD	P22662 macroptus ro
861	22	71.0	195	1	RM1_CAREL	O9wx94 clostridium	934	22	71.0	227	1	COX2_MACRO	P24988 microtus pe
862	22	71.0	198	1	MOBA_CLOPE	O60760 homo sapien	935	22	71.0	227	1	COX2_MOUSE	P00405 mus musculu
863	22	71.0	198	1	PCB2_HUMAN	O06383 mycobacteri	936	22	71.0	227	1	COX2_NYCP	O47675 nyctereutes
864	22	71.0	202	1	HPRT_MYCTU	P96794 mycobacteri	937	22	71.0	227	1	COX2_PHOYR	O00528 plica vitul
865	22	71.0	203	1	HPRT_MYCAV	O69537 mycobacteri	938	22	71.0	227	1	COX2_PROPA	O47676 pseudalopex
866	22	71.0	203	1	HPRT_MYCLE	P82125 sus scrofa	939	22	71.0	227	1	COX2_PSEBP	O47677 pseudalopex
867	22	71.0	206	1	AFR_PIG	P47276 mycoplasma	940	22	71.0	227	1	COX2_PSEBX	O47678 pseudalopex
868	22	71.0	206	1	UPP_MYCGE	P20473 tetrahymena	941	22	71.0	227	1	COX2_PSEBU	P80449 rattus norv
869	22	71.0	207	1	CB23_TETTH	Q9fm32 cyanidium c	942	22	71.0	227	1	COX2_RABIT	O61909 rhinoceros
870	22	71.0	208	1	EPFS_CYACA	O99388 mus musculu	943	22	71.0	227	1	COX2_RAT	P30691 sciurus car
871	22	71.0	208	1	Y24K_MOUSE	O60188 methanosarc	944	22	71.0	227	1	COX2_RH1UN	O47679 ovis aries
872	22	71.0	209	1	VAID_MITMA	P47165 saccharomyc	945	22	71.0	227	1	COX2_RH1UN	O47679 ovis aries
873	22	71.0	209	1	Y9JB_YEAST	P30113 schistosoma	946	22	71.0	227	1	COX2_SCTCA	O47679 ovis aries
874	22	71.0	211	1	GT28_SCHBO	P30114 schistosoma	947	22	71.0	227	1	COX2_SHEEP	O47679 ovis aries
875	22	71.0	211	1	GT28_SCHNA	P40337 homo sapien	948	22	71.0	227	1	COX2_SPEVE	O47679 ovis aries
876	22	71.0	213	1	VHL_HUMAN	P44331 mus musculu	949	22	71.0	227	1	COX2_STNCA	P50675 syncerus ca
877	22	71.0	215	1	I12A_MOUSE	P30661 callimico g	950	22	71.0	227	1	COX2_TARBA	P80443 tarsius ban
878	22	71.0	216	1	COX2_CALGO	P50664 african hor	951	22	71.0	227	1	COX2_TARBY	P37685 tragelaphus
879	22	71.0	217	1	VNS3_AHSV3	P09226 tetrahymena	952	22	71.0	227	1	COX2_TRAIM	O47680 vulpes macr
880	22	71.0	218	1	CB25_TETTH	P46661 cercocebus	953	22	71.0	227	1	COX2_VU1MA	O47681 vulpes vulp
881	22	71.0	219	1	I12A_CERTO	P29459 homo sapien	954	22	71.0	227	1	COX2_VU1VU	O47673 vulpes zerd
882	22	71.0	219	1	I12A_HUMAN	P48091 macaca mula	955	22	71.0	227	1	COX2_VU1ZE	P15530 mus musculu
883	22	71.0	219	1	I12A_MACMU	O86576 streptomyc	956	22	71.0	228	1	COX2_MOUSE	P30667 sus scrofa
884	22	71.0	219	1	Y1T1_STRCO	P36093 dictyosteli	957	22	71.0	228	1	COX2_PIG	O91025 streptomyc
885	22	71.0	220	1	NDKM_DICDI	P39313 escherichia	958	22	71.0	228	1	REPL_STRCO	P40259 homo sapien
886	22	71.0	220	1	YTFE_ECOLI	P37502 bacillus su	959	22	71.0	229	1	C79B_HUMAN	P00407 xenopus lae
887	22	71.0	220	1	YFBB_BACSV	O28231 cervus elap	960	22	71.0	230	1	COX2_XENLA	O37741 gadus morhu
888	22	71.0	221	1	I12A_BOVIN	O28267 canis famli	961	22	71.0	230	1	COX2_GADMO	P48171 oncorhynch
889	22	71.0	221	1	I12A_CEREL	O02743 felis silve	962	22	71.0	230	1	COX2_ONCMY	O37778 ornithorhyn
890	22	71.0	222	1	I12A_CANFA	O29053 sus scrofa	963	22	71.0	230	1	COX2_ORNAN	Q37677 salmo salar
891	22	71.0	222	1	I12A_FELCA	Q9sx96 equus cabal	964	22	71.0	230	1	COX2_SALSA	O79440 scylliorhnu
892	22	71.0	222	1	I12A_HORSE	O29053 sus scrofa	965	22	71.0	230	1	COX2_SCYCA	O32551 squallus aca
893	22	71.0	222	1	I12A_PIG	O54339 plasmodium	966	22	71.0	230	1	COX2_SOUAC	O15131 homo sapien
894	22	71.0	222	1	OS28_PLAGA	O57876 methanococ	967	22	71.0	231	1	EBP_HUMAN	P86024 alouatta pa
895	22	71.0	223	1	Y434_METLA	O61728 marmota mon	968	22	71.0	234	1	COX2_ALOPA	Q62z43 pyrodictulu
896	22	71.0	223	1	I12A_MARMO	P05467 kluyveromyc	969	22	71.0	234	1	PUR7_PYRAE	O39759 xenopus lae
897	22	71.0	225	1	YRFL_KLUTLA	Q9hg93 halobacteri	970	22	71.0	234	1	Y491_ARCFU	P29309 archaeoglob
898	22	71.0	225	1	URK_HALNI	P43543 saccharomyc	971	22	71.0	235	1	1433_XENLA	P11912 saccharomyc
899	22	71.0	225	1	YFGL_YEAST	Q37595 perognathus	972	22	71.0	237	1	SHR5_YEAST	P13131 staphylococ
900	22	71.0	226	1	COX2_PERFA	P50672 acornys wils	973	22	71.0	238	1	AGRA_STRAN	O93x67 neisseria m
901	22	71.0	227	1	COX2_ACOMI	Q37369 antilocapra	974	22	71.0	238	1	UBIG_NEIMA	O93x67 neisseria m
902	22	71.0	227	1	COX2_ANTAM	P50673 apodemus sy	975	22	71.0	238	1	HISA_MERTH	O68931 metanodact
903	22	71.0	227	1	COX2_APOSV	P42294 balaeopter	976	22	71.0	239	1	NOO2_PANDE	P29341 paracoccus
904	22	71.0	227	1	COX2_BALMU	P24986 balaeopter	977	22	71.0	240	1	LEF5_GVCL	P11721 cryptophleb
905	22	71.0	227	1	COX2_BALPH	Q37416 bison bison	978	22	71.0	240	1	PYRH_BACSV	O31749 bacillus su
906	22	71.0	227	1	COX2_BISBI	P50674 bos javanic	979	22	71.0	240	1	Y136_ARATH	O98h63 rhizobium l
907	22	71.0	227	1	COX2_BOSJA	O37419 boselaphus	980	22	71.0	241	1	FLGF_RHILLO	
908	22	71.0	227	1	COX2_BOSIR	P00404 bos taurus	981	22	71.0	241	1	GTS_MUSDO	
909	22	71.0	227	1	COX2_BOVIN		982	22	71.0	241	1		

ALIGNMENTS

PSMR_METUA	STANDARD	PRT	430 AA
983	22	71.0	241
984	22	71.0	244
985	22	71.0	244
986	22	71.0	245
987	22	71.0	245
988	22	71.0	245
989	22	71.0	245
990	22	71.0	245
991	22	71.0	245
992	22	71.0	245
993	22	71.0	245
994	22	71.0	245
995	22	71.0	245
996	22	71.0	245
997	22	71.0	245
998	22	71.0	245
999	22	71.0	246
1000	22	71.0	246

19224 ratu norv
P2213 myrocheium
060471 cavia porce
P29378 bos taurus
P1946 homo sapien
P35213 mus musculu
P27348 homo sapien
P35216 mus musculu
P29312 homo sapien
P35215 mus musculu
P29361 ovis aries
081512 anabaena sp
019010 equus cabal
P29359 bos taurus
P35214 homo sapien

IS 80 DEGREES CELSIUS.
-1- SIMILARITY: BELONGS TO THE AAA FAMILY OF ATPASES.
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).
EMBL: U67559; AAB99179.1; -
TIGR: M11176; -
InterPro: IPR003593; AAA_ATPase.
InterPro: IPR003959; AAA_ATPase_centrl.
InterPro: IPR003960; AAA_sub.
Pfam: PF00004; AAA; 1.
SMART: SM00382; AAA; 1.
PROSITE: PS00674; AAA; 1.
TIGRATMS: TIGR01242; 265p45; 1.
Kw Proteasome; ATP-binding; Complete proteome.
NP_BIND 211 218 ATP (POTENTIAL).
SQ SEQUENCE 430 AA: 48689 MW: 3FD2E94A68D463DD CRC64;

Query Match 93.5%; Score 29; DB 1; Length 430;
Best Local Similarity 83.3%; Pred. NO. 45;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDED 6
DB 167 VREDED 172

RESULT 1
PSMR_METUA STANDARD; PRT; 430 AA.
AC 058576; 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN PAN OR M1176.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D., Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D., Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I., Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A., Scott J.L., Geoghegan N.S.M., Feldman J.F., Fuhrmann J.L., Nguyen D., Uuterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C., Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M., Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.";
RL Science 273:1058-1073(1996).
[2]
RN CHARACTERIZATION.
RX MEDLINE=20158885; PubMed=10692374;
RA Wilson H.L., Ou M.S., Aldrich H.C., Maupin-Furlow J.;
RT "Biochemical and physical properties of the Methanococcus jannaschii 20S proteasome and PAN, a homolog of the ATPase (Rpt) subunits of the eucaryal 26S proteasome.";
RL J. Bacteriol. 182:1680-1692(2000).
[3]
RN CHARACTERIZATION.
RX MEDLINE=99403035; PubMed=10473546;
RA Zwickl P., Ng D., Woo K.M., Klenk H.-P., Goldberg A.L.;
RT "An archaeobacterial ATPase, homologous to ATPases in the eukaryotic 26 S proteasome, activates protein breakdown by 20 S proteasomes.";
RL J. Biol. Chem. 274:26008-26014(1999).
CC -1- FUNCTION: Required for the ATP- or CTP-dependent degradation of proteins, but not small peptides, by the 20S proteasome.
CC -1- COFACTOR: MAGNESIUM.
CC -1- ENZYME REGULATION: INHIBITED BY EDTA, NEM AND POMS.
CC -1- SUBUNIT: HOMOHexamER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- MISCELLANEOUS: ITS OPTIMUM PH IS 7-8 AND THE OPTIMUM TEMPERATURE

RESULT 2
WIS4_SCHPO STANDARD; PRT; 1401 AA.
AC 014289; P87062; Q92384;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE MAP kinase kinase kinase wis4 (EC 2.7.1.1.) (MAP kinase kinase kinase DE waki) (MAP kinase kinase kinase wis4 EC 2.7.1.1.) (MAP kinase kinase kinase wis4 OR WAKI OR WIKI OR SPAC9G1.02).
GN Schizosaccharomyces pombe (fission yeast).
OS Schizosaccharomycetes; Fungi; Ascomycota; Schizosaccharomycetes;
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=98026892; PubMed=9321395;
RA Samejima I., Mackie S., Fantes P.A.;
RT "Multiple modes of activation of the stress-responsive MAP kinase pathway in fission yeast.";
RL EMBO J. 16:6162-6170(1997).
[2]
RN SEQUENCE FROM N.A.
RP STRAIN-972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A., Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M., Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A., Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G., Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K., James K., Jones L., Jones M., Leather S., McDonald S., McLean J., Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C., Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E., Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S., Skelton J., Simmonds M., Squares R., Squares K., Stevens K., Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S., Woodward J., Volckaert G., Aert R., Robben J., Grymponerz B., Welfens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,

RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purrelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelure V., Mottier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Horst S.M.,
 RA Lucas M., Kochet M., Gallardin C., Tallada V.A., Garzon A., Thode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.,
 RA "The genome sequence of *Schizosaccharomyces pombe*,"
 RT Nature 415:871-880(2002).
 RN [3]
 RP SEQUENCE OF 96-1401 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97282620; PubMed=9136929;
 RA Miller J.C., Wilkinson M., Buck V., Morgan B., Makino K.,
 RA Shier J.B.A.;
 RT "The Mcs4 response regulator coordinately controls the stress-
 RT activated Wsk1-Wsk1-Styl MAP kinase pathway and fission yeast cell
 RT cycle.";
 RL Genes Dev. 11:1008-1022(1997).
 RN [4]
 RP SEQUENCE OF 457-543 FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=97331762; PubMed=9188094;
 RA Shiozaki K., Shiozaki M., Russell P.,
 RA "Mcs4 mitotic catastrophe suppressor regulates the fission yeast cell
 RT cycle through the Wsk1-Wsk1-Spc1 kinase cascade.";
 RT Mol. Biol. Cell 8:409-419(1997).
 CC -1- FUNCTION: INVOLVED IN A SIGNAL TRANSDUCTION PATHWAY THAT IS
 CC ACTIVATED IN UNDER CONDITIONS OF HEAT SHOCK, OXIDATIVE STRESS OR
 CC LIMITED NUTRITION. UNLIKE WSK1, IT IS NOT ACTIVATED BY CHANGES IN
 CC THE OSMOLARITY OF THE EXTRACELLULAR ENVIRONMENT. ACTIVATES THE
 CC WSK1 MAP KINASE KINASE BY PHOSPHORYLATION.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC MAP KINASE KINASE SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: 298763; CA011500.1; -
 DR EMBL: Y07763; CA069030.1; ALU-INT1.
 DR EMBL: Y11969; CA02718.1; -
 DR EMBL: U81521; AAB39762.1; -
 DR HSSP: P24941; ICRP.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR002290; Ser_thr_pkinase.
 DR Pfam: PF00069; pkinase.1.
 DR ProDom: PD000001; Euk_pkinase.1.
 DR SMART: SM00220; S_TKc.1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM.1.
 DR PROSITE: PS00108; PROTEIN_KINASE_ST.1.
 KW Transferrase: Serine/threonine-protein kinase; ATP-binding.
 FT DOMAIN 1037 1306 PROTEIN KINASE.
 FT NP_BIND 1043 1051 ATP (BY SIMILARITY).
 FT BINDING 1066 1066 ATP (BY SIMILARITY).
 FT ACT_SITE 1161 1161 BY SIMILARITY.
 FT CONFLICT 484 485 RL -> SP (IN REF. 3).
 FT CONFLICT 1346 1346 D -> V (IN REF. 1).
 SO SEQUENCE 1401 AA; 160537 MW; 529A989D2B627F3D CRC64;

Query Match 93.5%; Score 29; DB 1; Length 1401;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDI 6

DB 1194 VREDI 1199
 |||||
 RESULT 3
 ID RPOC_TREPA STANDARD; PRI: 1416 AA.
 AC 083270;
 DT 15-DEC-1998 (rel. 37, Created)
 DT 15-DEC-1998 (rel. 37, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE DNA-directed RNA polymerase beta' chain (EC 2.7.7.6) (Transcriptase
 DE beta' chain) (RNA polymerase beta' subunit).
 GN RPOC OR TP0242.
 OS Treponema pallidum.
 OC Bacteria; Spirochaetales; Spirochaetaceae; Treponema.
 OX NCBI_Taxid=160;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Nichols;
 RX MEDLINE=96832770; PubMed=9665876;
 RA Fraser C.M., Norris S.J., Weinstock G.M., White O., Sutton G.G.,
 RA Dodson R., Gwin M., Hickey E.K., Clayton R., Ketchum K.A.,
 RA Sodergren E., Hardam J.M., McLeod M.P., Salzberg S., Peterson J.,
 RA Khalak H., Richardson D., Howell J.K., Chidambaram M., Uitterback L.,
 RA McDonald L., Artlich P., Bowman C., Cotton M.D., Fujii C., Garland S.,
 RA Hatch B., Horst K., Roberts K., Sandusky M., Weidman J., Smith H.O.,
 RA Venter J.C.;
 RT "Complete genome sequence of *Treponema pallidum*, the syphilis
 RT spirochete.";
 RT Science 281:375-388(1998).
 RL -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 CC SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
 CC [RNA(N)].
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 CC ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 CC BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----
 DR EMBL: AE001205; AAC65230.1; -
 DR HSSP: O9KW06; IHOM.
 DR TIGR: TP0242.
 DR InterPro: IPR000722; RNA_POL_A.
 DR InterPro: IPR002879; RNA_POL_A2.
 DR Pfam: PF00623; RNA_POL_A.1.
 DR Pfam: PF01854; RNA_POL_A2.2.
 KW Transferrase: DNA-directed RNA polymerase; Transcription;
 KW complete proteome.
 SO SEQUENCE 1416 AA; 159789 MW; 78D0EB6A1CAF7C9C CRC64;

Query Match 93.5%; Score 29; DB 1; Length 1416;
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDI 6

DB 1052 VREDI 1057

RESULT 4
 ID CHIS_BACSU STANDARD; PRI: 277 AA.
 AC 007921;

DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Chitosanase precursor (EC 3.2.1.132).
 GN CSN.
 OS *Bacillus subtilis*.
 CC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97266548; PubMed=9141695;
 RA Paro V., San Roman M., Gallardo I., Purnelle B., Bolotin A.,
 RA Sorokin A., Mellado R.P.,
 RT "A 2391 bp region of the *Bacillus subtilis* genome comprising genes
 RT located upstream and downstream of the *lev* operon."
 RL Microbiology 143:1321-1326 (1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=97453479; PubMed=9308178;
 RA Sorokin A., Bolotin A., Purnelle B., Hilbert H., Lauber J.,
 RA Duesterhoeft A., Ehrlich S.D.,
 RT "Sequence of the *Bacillus subtilis* genome region in the vicinity of
 RT the *lev* operon reveals two new extracytoplasmic function RNA
 RT polymerase sigma factors *SigV* and *SigZ*."
 RL Microbiology 143:2939-2943 (1997).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Azevedo V., Berrero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Bortiss R., Bourschi L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codan J.J., Conerton I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Duesterhoeft A., Ehrlich S.D., Emmerson P.T.,
 RA Enlian K.D., Errington J., Fabret C., Ferrari E., Follger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Gallizi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Hachez J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holstappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Klierer-Blanchard M., Klein C.,
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue V.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
 RA Paro V., Pohl T.M., Portelle D., Portwolk S., Prescott A.M.,
 RA Priescan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roeha E., Roche B., Rose M., Sadele Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska E., Seror S.J., Serror P., Shin B.S., Soldo B.,
 RA Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tanakoshi A., Tanaka T., Terpetra P., Tognoni K.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
 RA Viati A., Wandut R., Wedler E., Wedler H., Weitenegger T.,
 RA Whittier P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.,
 RT "The complete genome sequence of the Gram-positive bacterium *Bacillus*
 RT *subtilis*."
 RL Nature 390:249-256 (1997).
 CC -1- FUNCTION: AID IN THE DEFENSE TO INVADING FUNGAL PATHOGENS BY
 CC DEGRADING THEIR CELL WALL CHITOSAN (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: Endohydrolysis of beta-1,4-linkages between N-
 CC acetyl-D-glucosamine and D-glucosamine residues in a partly
 CC acetylated chitosan.
 CC -1- SUBCELLULAR LOCATION: Secreted (By similarity).
 CC -1- SIMILARITY: BELONGS TO FAMILY 46 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X92868; CA63455.1;
 CC DR EMBL: U93875; AAB80882.1;
 CC DR EMBL: 299117; CAB14630.1;
 CC DR WSSP: P33665; ICHK.
 CC DR SUBLIST: BG11923; csn.
 CC DR InterPro: IPR00400; GH_46.
 CC DR Pfam: PF01374; Glyco_hydro_46; 1.
 CC DR PROSITE: PS60000; CHITOSANASE_46_80; 1.
 CC KM Hydrolyase; Glycosidase; Signal; Complete proteome.
 CC FT SIGNAL 1 35
 CC FT CHAIN 36 277
 CC FT ACT_SITE 54 54
 CC FT ACT_SITE 70 70
 CC FT SIGNAL 1 35
 CC FT ACT_SITE 54 54
 CC FT ACT_SITE 70 70
 CC SQ SEQUENCE 277 AA: 31496 MW: 8C17C156AE27F781 CRC64;
 CC -----
 CC Query Match 90.3%; Score 28; DB 1; Length 277;
 CC Best Local Similarity 83.3%; Pred. No. 47;
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 VRYEDL 6
 CC Db 222 VRYDDL 227
 CC -----
 CC RESULT 5
 CC CAC3_STRPN STANDARD: PRT: 306 AA.
 CC ID CAC3_STRPN
 CC AC 054800; Q54612;
 CC DT 01-NOV-1997 (Rel. 35, Created)
 CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE UTP--glucose-1-phosphate uridylyltransferase (EC 2.7.7.9) (UDP-glucose
 CC DE pyrophosphorylase) (UDPgc) (Alpha-D-glucosyl-1-phosphate
 CC DE uridylyltransferase) (uridine diphosphoglucose pyrophosphorylase).
 CC GN CAP3 OR CP83U.
 CC OS Streptococcus pneumoniae.
 CC CC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
 CC CC Streptococcus.
 CC OX NCBI_TaxID=1313;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=406 / Type 3;
 CC RX MEDLINE=96144241; PubMed=8566758;
 CC RA Arcubietta C., Garcia E., Lopez R.;
 RT "Sequence and transcriptional analysis of a DNA region involved in
 RT the production of capsular polysaccharide in *Streptococcus pneumoniae*
 RT type 3."
 RL Gene 167:1-7 (1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=M02 / Serotype 3;
 RX MEDLINE=95173611; PubMed=7869055;
 RA Dillard J.P., Vandersa M.W., Vother J.;
 RT "Characterization of the cassette containing genes for type 3
 RT capsular polysaccharide biosynthesis in *Streptococcus pneumoniae*."
 RL J. Exp. Med. 181:973-983 (1995).
 CC -1- CATALYTIC ACTIVITY: UTP + alpha-D-glucose 1-phosphate =
 CC diphosphate + UDP-glucose.
 CC -1- PATHWAY: PRODUCTION OF CAPSULAR POLYSACCHARIDE.
 CC -1- SIMILARITY: BELONGS TO THE PROKARYOTIC UDPgc FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

```

CC -----
DR EMBL: Z47210; CAA87405.1; -
DR EMBL: U15171; AAC43313.1; -
DR InterPro: IPR001825; NTP_transferase.
DR Pfam: PF00483; NTP_transferase; 1.
DR TIGRfams: TIGR01099; galU; 1.
KW TRANSFERASE; Kinase; Nucleotidyltransferase.
FT CONFLICT 204 204 N->H (IN REF. 2).
SQ SEQUENCE 306 AA; 33919 MW; F8E243D00F47EBA8 CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DA 1; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYEDL 6
Db 165 VYEDV 170

RESULT 6
ASRC_SALTY STANDARD; PRT; 337 AA.
ID ASRC_SALTY
AC P26476;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Anaerobic sulfite reductase subunit C (EC 1.8.1.-).
GN ASRC OR STM2550 OR STY2796.
OS Salmonella typhimurium, and
OC Salmonella typhi.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhimurium; STRAIN-EB303;
RA MEDLINE-91139599; PubMed-1704886;
RA Huang C.J., Barrett E.L.;
RT *Sequence analysis and expression of the Salmonella typhimurium asr
RT operon encoding production of hydrogen sulfide from sulfite.*;
RL J. Bacteriol. 173:1544-1553(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhimurium; STRAIN-LT2 / SGSC1412 / ATCC 700720;
RA MEDLINE-21534948; PubMed-11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT *Complete genome sequence of Salmonella enterica serovar Typhimurium
RT LT2.*;
RL Nature 413:852-856(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES-S. typhi; STRAIN-CT18;
RA MEDLINE-21534947; PubMed-11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Main J.,
RA Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
RA Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
RA Croft A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA Felwell I., Hamlin N., Haque A., Hien T.T., Holroyd S., Jags K.,
RA Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA Quail M., Rutherford K., Simmonds M., Skelton J., Stevens K.,
RA Whitehead S., Barrall B.G.;
RT *Complete genome sequence of a multiple drug resistant Salmonella
RT enterica serovar Typhi CT18.*;
RL Nature 413:848-852(2001).
CC -1- FUNCTION: THIS ENZYME CATALYZES THE HYDROGEN SULFIDE PRODUCTION
CC FROM SULFITE. IT IS STRICTLY ANAEROBIC. IT IS REGULATED BY
CC ELECTRON ACCEPTORS RATHER THAN BY CYSTEINE.
CC -1- CATALYTIC ACTIVITY: Hydrogen sulfide + 3 NAD(+) + 3 H(2)O =
CC sulfite + 3 NADH.

```

```

CC -----
CC -1- COFACTOR: THIS PROTEIN CONTAINS TWO SIROHEME AND EIGHT 4FE-4S
CC IRON-SULFUR CENTERS AS PROSTHETIC GROUPS.
CC -1- PATHWAY: Sulfite reduction.
CC -1- SUBUNIT: THE ANAEROBIC SULFITE REDUCTASE SEEMS TO CONSIST OF THREE
CC SUBUNITS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- INDUCTION: BY SULFITE.
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN IS A 4FE-4S/SIROHEME DOMAIN
CC FOUND IN NITRATE REDUCTASES (EC 1.6.6.4 AND EC 1.7.7.1) AND
CC SULFITE REDUCTASES (EC 1.8.1.2 AND EC 1.8.7.1).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M57706; AAA99277.1; -
DR EMBL: AE008816; AAL21444.1; -
DR EMBL: AL627275; CAD02753.1; -
DR PIR: C38453; C38453.
DR StyGene; SG10023; asrC.
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR InterPro: IPR000660; Ntr_sir.
DR InterPro: IPR005117; Ntr_sir_fer.
DR Pfam: PF00037; fer4_2.
DR Pfam: PF01077; Ntr_sir; 1.
DR Pfam: PF03460; Ntr_sir_fer; 1.
DR PROSITE; PS00198; 4FE4S_FERREDOXIN; 1.
DR PROSITE; PS00365; Ntr_sir; 1.
KW Oxidoreductase; Electron transport; Heme; NAD; Iron-sulfur; 4Fe-4S;
KW Complete proteome.
FT METAL 115 115 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 121 121 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 153 153 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 157 157 IRON-SULFUR OR HEME IRON (POTENTIAL).
FT METAL 180 180 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 183 183 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 186 186 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 190 190 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 212 212 IRON-SULFUR 1 (4FE-4S) (BY SIMILARITY).
FT METAL 215 215 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 218 218 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT METAL 222 222 IRON-SULFUR 2 (4FE-4S) (BY SIMILARITY).
FT CONFLICT 247 250 TPRV -> SGAL (IN REF. 1).
SQ SEQUENCE 337 AA; 37290 MW; 183FE2F69334CBF CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DA 1; Length 337;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYEDL 6
Db 69 IYEDV 74

RESULT 7
YMG7_MYCTU
ID YMG7_MYCTU STANDARD; PRT; 388 AA.
AC O50695;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein RV2267C.
GN RV2267C OR MT2329 OR MYCY339.43.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;

```


[1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=H37RV;
 RC MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Teketa C.,
 RA Baccok K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skellern S., Squares S., Squares R.,
 RA Stulton J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J.F., Deboy R., Dodson R., Gwin M.L., Haft D., Hickey L.,
 RA Kolonay J.F., Nelson W.C., Umayan L.A., Emdolova M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Kouri H., Gill J., Mikula A.,
 RA Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC
 CC EMBL; Z77163; CAB00968.1; -;
 DR EMBL; AE007076; AAK46511.1; -;
 DR TIGR; MT2329; -;
 DR Tuberculist; RV2267c; -;
 DR InterPro: IPR000863; Sulfoltransferase.
 DR Pfam; PF00685; Sulfoltransfer: 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 388 AA; 46062 MW; 5DED0263275A9B24 CRC64;

 Query Match 90.3%; Score 28; DB 1; Length 388;
 Best Local Similarity 83.3%; Pred. No. 68;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 VRYEDL 6
 Db 302 LRYEDL 307

 RESULT 8
 MOEA_ANASP STANDARD: PRT; 436 AA.
 ID MOEA_ANASP
 AC Q44243;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Molybdopterin biosynthesis protein moea.
 GN MOEA OR ALB5136.
 OS Anabaena sp. (Strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96272265; PubMed=8682795;
 RA Ramasamy K.S., Endley S., Golden J.W.;
 RT "Nitrate reductase activity and heterocyst suppression on nitrate in
 RL Anabaena sp. strain PCC 7120 require moea.";
 J. Bacteriol. 178:3893-3898(1996).
 [2]

RP SEQUENCE FROM N.A.
 RX MEDLINE=21595285; PubMed=11759840;
 RA Kaneo T., Nakamura Y., Wolk C.P., Sasamoto S.,
 RA Watanabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Kuraki A.,
 RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT Complete genomic sequence of the filamentous nitrogen-fixing
 RT cyanobacterium Anabaena sp. strain PCC 7120.";
 RL Nucleic Acids Res. 29:205-213(2001).
 CC -1- FUNCTION: INVOLVED IN THE BIOSYNTHESIS OF A DEMOLYBDO-COFACTOR
 CC (MOLYBDOPEPTIN), NECESSARY FOR MOLYBDO-ENZYMES (BY SIMILARITY).
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
 CC -1- SIMILARITY: NO THE C-TERMINAL OF CINNAMON/GERANYRIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 CC
 CC EMBL; U34309; AAC44505.1; -;
 DR EMBL; AP003598; BAB76835.1; -;
 DR InterPro: IPR001453; MOCF_biosynth.
 DR InterPro: IPR005111; Moea.C.
 DR InterPro: IPR005110; Moea.N.
 DR Pfam; PF00994; MOCF_biosynth: 1.
 DR Pfam; PF03453; Moea.N. 1.
 DR Pfam; PF03454; Moea.C. 1.
 DR Pfam; PF002460; MOCF_biosynth: 1.
 DR ProDom; PD002460; MOCF_biosynth: 1.
 DR TIGR; TIGR00177; molyb-syn: 1.
 DR PROSITE; PS01079; MOCF BIOSYNTHESIS 2; 1.
 KW Molybdenum cofactor biosynthesis; Complete proteome.
 SQ SEQUENCE 436 AA; 47332 MW; 65D732E44336561 CRC64;

 Query Match 90.3%; Score 28; DB 1; Length 436;
 Best Local Similarity 83.3%; Pred. No. 77;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

 QY 1 VRYEDL 6
 Db 84 VRYEDV 89

 RESULT 9
 C6ST_CHICK STANDARD: PRT; 458 AA.
 ID C6ST_CHICK
 AC Q92179;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Chondroitin 6-sulfoltransferase (EC 2.8.2.17) (C6ST).
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-44; 46-52 AND 186-192.
 RC STRAIN=White Leghorn; TISSUE=Embryonic chondrocytes;
 RX MEDLINE=95355490; PubMed=7629189;
 RA Fukuta K., Uchimura K., Nakashima K., Kato M., Kimata K.,
 RA Shinomura T., Habuchi O.;
 RT "Molecular cloning and expression of chick chondrocyte chondroitin 6-
 RT sulfoltransferase.";
 RT J. Biol. Chem. 270:18575-18580(1995).
 CC -1- FUNCTION: CATALYZES THE SULFATION OF CHONDROITIN AND KERATAN
 CC SULFATE.
 CC -1- CATALYTIC ACTIVITY: 3'-phosphoadenylylsulfate + chondroitin =
 CC adenosine 3',5'-bisphosphate + chondroitin 6'-sulfate.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. GOLGI MEMBRANE

(BY SIMILARITY).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D49915; BAA08655.1; -
KM Transferrase; Transmembrane; Signal-anchor; Glycoprotein; Golgi stack.
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 37 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 38 458 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 95 95 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 236 236 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 443 443 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 458 AA; 52253 MW; C9A3B7D0A5086F0C CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 458;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 348 VREEDV 353
|||||

RESULT 10
AHPE-STAAU STANDARD; PRT; 507 AA.
ID AHPE-STAAU 005204;
AC 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DR 15-DEC-1998 (Rel. 37, Last annotation update)
DE ALK1 hydroperoxide reductase subunit F (EC 1.6.4.-).
GN AHPE.
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RM4220.
RA Jones E.C., Francis K.P., Stewart G.S.A.B.;
RL Submitted (MAR-1997) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: SERVES TO PROTECT THE CELL AGAINST DNA DAMAGE BY ALKYL
CC HYDROPEROXIDES. IT CAN USE EITHER NADH OR NADPH AS ELECTRON DONOR
CC FOR DIRECT REDUCTION OF REDOX DYES OR OF ALKYL HYDROPEROXIDES WHEN
CC COMBINED WITH THE AHPC PROTEIN.
CC -1- COFACTOR: FAD.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- MISCELLANEOUS: THE ACTIVE SITE IS A REDOX-ACTIVE DISULFIDE BOND.
CC -1- SIMILARITY: BELONGS TO THE PYRIDINE NUCLEOTIDE-DISULFIDE
CC OXIDOREDUCTASES CLASS-II.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U92441; AAB51152.1; -
DR HSSP: P09625; 1TRB.
DR InterPro: IPR001327; FAD_Pyr_redox.
DR InterPro: IPR000103; Pyridine_redox_2.
DR Pfam: PF00070; Pyr_redox; 1.
DR PRINTS: PR00368; FADPNR.

DR ProDom: PD000139; FAD_Pyr_redox; 1.
DR PROSITE: PS00573; PYRIDINE_REDOX_2; 1.
KM Redox-active center; Oxidoreductase; NADP; NAD; Flavoprotein; FAD.
FT NP_BIND 207 222 FAD (ADP PART) (BY SIMILARITY).
FT DISULFID 335 338 REDOX-ACTIVE (BY SIMILARITY).
FT NP_BIND 347 361 NAD(P) (BY SIMILARITY).
FT NP_BIND 467 477 FAD (FLAVIN PART) (BY SIMILARITY).
SQ SEQUENCE 507 AA; 54720 MW; EDED04232731D21E CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 507;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
DB 416 IREEDM 421
|||||

RESULT 11
M3K3-MOUSE STANDARD; PRT; 626 AA.
ID M3K3-MOUSE 061084;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DR 30-May-2000 (Rel. 39, Last annotation update)
DE Mitogen-activated protein kinase kinase 3 (EC 2.7.1.-)
DE (MAPK/ERK kinase kinase 3) (MEK kinase 3) (MEK 3).
GN MAPK3 OR MEK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain.
RX MEDLINE=96214986; PubMed=8621389.
RA Blank J.L., Gervins P., Elliott E.M., Sather S., Johnson G.L.;
RT "Molecular cloning of mitogen-activated protein/ERK kinase kinases
RT (MEK) 2 and 3. Regulation of sequential phosphorylation pathways
RT involving mitogen-activated protein kinase and c-Jun kinase.";
RL J. Biol. Chem. 271:5361-5368(1996).
CC -1- FUNCTION: PREFERENTIALLY ACTIVATES P42/44 (ERK2/ERK) MAP KINASES.
CC -1- SIMILARITY: BELONGS TO THE SPR/THR FAMILY OF PROTEIN KINASES.
CC MAP KINASE KINASE KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: U43187; AAB03535.1; -
DR HSSP: Q00534; 1B18.
DR MGD: MGI:1346874; MapK3.
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_Chr_Pkinase.
DR Pfam: PF00069; Pkinase; 1.
DR ProDom: PD000001; Euk_Pkinase; 1.
DR SMART: SM00220; S_TKC; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; FALSE NEG.
DR PROSITE: PS00108; PROTEIN_KINASE_ST; FALSE NEG.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KM Transferrase; Serine/threonine-protein kinase; ATP-binding.
FT DOMAIN 362 622 PROTEIN KINASE.
FT NP_BIND 368 375 ATP (BY SIMILARITY).
FT BINDING 391 391 ATP (BY SIMILARITY).
FT ACT_SITE 489 489 BY SIMILARITY.
SQ SEQUENCE 626 AA; 70775 MW; 00EF2443C9556E0B CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 1; Length 626;
Matches 8; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYEDL 6
|||||
Db 64 VREYEDV 69

RESULT 12

YJEP-ECOLI STANDARD; PRT: 1107 AA.
AC P39285; P76798; Rel. 31, Created)
DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein yjep precursor.
GN YJEP OR B4159.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN SEQUENCE FROM N.A.
RP STRAIN-K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RX Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE OF 1-27 FROM N.A.
RC SPRAIN-K12;
RX MEDLINE=88298809; PubMed=3042771;
RA Li Q.X., Dowhan W.;
RT "Structural characterization of Escherichia coli phosphatidylserine
decarboxylase.";
RL J. Biol. Chem. 263:11516-11522(1988).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE UPF0003 FAMILY. STRONG, TO H. INFLUENZA
CC A10195.1.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U14003; AAA97058.1; -
DR EMBL: AE000488; AAC77119.1; -
DR EMBL: J03916; AAA83897.1; ALT_INT.
DR EcoGene; EG12478; yjep.
DR InterPro; IPR01880; MSion_channel.
DR Pfam; PF00924; MS_channel.1.
DR PROSITE; PS01246; UPF0003.1.
KW Hypothetical protein; Transmembrane; Signal; Complete proteome.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 1107 HYPOTHELTICAL PROTEIN YJEP.
FT TRANSMEM 467 487 POTENTIAL.
FT TRANSMEM 551 571 POTENTIAL.
FT TRANSMEM 600 620 POTENTIAL.
FT TRANSMEM 628 648 POTENTIAL.

FT TRANSMEM 674 694 POTENTIAL.
FT TRANSMEM 698 718 POTENTIAL.
FT TRANSMEM 785 805 POTENTIAL.
FT TRANSMEM 828 848 POTENTIAL.
FT TRANSMEM 875 895 POTENTIAL.
FT TRANSMEM 910 930 POTENTIAL.
FT TRANSMEM 915 1015 POTENTIAL.
FT CONFLICT 1015 1015 R -> A (IN REF. 1).
SQ SEQUENCE 1107 AA; 123967 MW; 5F52A293B9532B CRC64;

Query Match 90.3%; Score 28; DB 1; Length 1107;
Best Local Similarity 83.3%; Pred. No. 2.1e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYEDL 6
|||||
Db 348 VREYEDL 353

RESULT 13

RRPL_PI3H4 STANDARD; PRT: 2233 AA.
AC P12577;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE RNA polymerase beta subunit (Ec 2.7.7.48) (large structural protein).
GN L.
OS Human parainfluenza 3 virus (strain NIH 47885).
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Respirovirus.
OX NCBI_TaxID=11217;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88306242; PubMed=2841798;
RA Gallinski M.S., Mink M.A., Pons M.W.;
RT "Molecular cloning and sequence analysis of the human parainfluenza 3
RT virus gene encoding the L protein.";
RL Virology 165:499-510(1988).
RN [2]
RP SEQUENCE OF 1-26 FROM N.A.
RX MEDLINE=88032139; PubMed=2822598;
RA Storey D.G., Cole M.-J., Dimock K., Kang C.Y.;
RT "Nucleotide sequence of the coding and flanking regions of the human
RT parainfluenza virus 3 hemagglutinin-neuraminidase gene: comparison
RT with other paramyxoviruses.";
RL Intervirology 27:69-80(1987).
CC -1- FUNCTION: PROBABLE COMPONENT OF THE ACTIVE POLYMERASE. IT MAY
CC FUNCTION IN RNA SYNTHESIS, CAPING, METHYLATION AND POLY(A)
CC SYNTHESIS OF NEWLY SYNTHESIZED VIRAL MRNAS. RNA EDITING OF THE P
CC GENE TRANSCRIPT, AND PROTEIN KINASE ACTIVITY.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -1- SIMILARITY: WITH L PROTEIN OF OTHER PARAMYXOVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M21649; AAA46854.1; -
DR EMBL: M20402; AAA46857.1; -
DR PIR: A29246; ZLN2P3.
DR InterPro; IPR001016; Viral_RNA_pol_L.
DR Pfam; PF00946; Paramyx_RNA_pol.1.
KW Transcriptase; RNA-directed RNA polymerase.
FT CONFLICT 26 26 K -> R (IN REF. 2).
SQ SEQUENCE 2233 AA; 255802 MW; DBB8B5DE74B4638 CRC64;

Query Match 90.3%; Score 28; DB 1; Length 2233;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Best Local Similarity 66.7%; Pred. No. 4.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 1126 IRYEDM 1131

RESULT 14

UREE_CONGL STANDARD; PRT; 157 AA.
ID UREH3; Q9L419;
AC 16-OCT-2001 (Rel. 40, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
GN UREase accessory protein uree.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13869;
RX MEDLINE=2126277; PubMed=11328647;
RA Puskas L.G., Inui M., Yukawa H.;
RT "Structure of the urease operon of Corynebacterium glutamicum.";
RL DNA Seq. 11:383-394(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
RX MEDLINE=20389622; PubMed=10930756;
RA Nolden L., Beckers G., Moeckel B., Pfeifferle W., Nampoothiri K.M.,
RA Kraemer R., Burkowski A.;
RT "Urease of Corynebacterium glutamicum: organisation of corresponding
genes and investigation of activity.";
RL FEMS Microbiol. Lett. 189:305-310(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT Submitted (May-2002) to the EMBL/Genbank/DSJ databases.
RL -1- FUNCTION: INVOLVED IN UREASE METALLOCENTER ASSEMBLY. BINDS NICKEL.
CC PROBABLY FUNCTIONS AS A NICKEL DONOR DURING METALLOCENTER ASSEMBLY
(BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE UREASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AB029154; BAA8555.1; -;
DR EMBL: AJ251883; CAB81938.1; -;
DR EMBL: AP005274; BAB97480.1; -;
DR InterPro: IPR004029; UREase.
DR Pfam: PF02814; uree; 1.
KW Nickel.
FT CONFLICT 42 42 P -> S (IN REF. 1).
FT CONFLICT 52 57 NEGHPI -> ITGFS (IN REF. 1).
FT CONFLICT 69 69 K -> R (IN REF. 1).
FT CONFLICT 133 157 DEHGLEYORTLEVPIPRHSGHTH -> G (IN REF. 1).
SQ SEQUENCE 157 AA; 17644 MW; 4BBAA3381ACDSC21 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 157;
Best Local Similarity 83.3%; Pred. No. 42;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 25 VREEDM 30

RESULT 15

MLE_TODPA STANDARD; PRT; 159 AA.
ID MLE_TODPA
AC P05945;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-NOV-1988 (Rel. 09, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Myosin catalytic light chain LC-1, mantle muscle.
OS Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Teuthoidea;
OC Oegopsida; Ommastrephidae; Todarodes.
OX NCBI_TaxID=6637;
RN [1]
RP SEQUENCE
RX MEDLINE=87076038; PubMed=3790251;
RA Watanabe B., Malta T., Konno K., Matsuda G.;
RT "Amino-acid sequence of LC-1 light chain of squid mantle muscle
myosin.";
RL Biol. Chem. Hoppe-Seyler 367:1025-1032(1986).
CC -1- FUNCTION: IN MOLLUSCAN MUSCLE, CALCIUM REGULATION IS ASSOCIATED
CC WITH MYOSIN RATHER THAN WITH ACTIN. MUSCLE MYOSIN CONTAINS TWO
CC TYPES OF LIGHT CHAINS: THE CATALYTIC LIGHT CHAIN, ESSENTIAL FOR
CC ATPASE ACTIVITY, AND THE REGULATORY LIGHT CHAIN, A CALCIUM-BINDING
CC PROTEIN RESPONSIBLE FOR CA(++) DEPENDENT BINDING AND CA(++)
CC DEPENDENT MG-ATPASE ACTIVITY.
CC -1- SIMILARITY: TO OTHER EF-HAND CALCIUM BINDING PROTEINS, BUT THIS
CC PROTEIN DOES NOT BIND CALCIUM.
DR PIR: A25571; A25571.
DR HSPD: F07291; IMDC.
DR InterPro: IPR002048; EF-hand.
DR Pfam: PF00036; efhand; 1.
DR Prodom: PD00012; EF-hand; 2.
KW Myosin; Muscle protein.
FT MOD_RES 1
SQ SEQUENCE 159 AA; 18059 MW; 0F60334A3A9711410 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 159;
Best Local Similarity 66.7%; Pred. No. 43;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
Db 140 IRYEDL 145

RESULT 16

GP30_BPSP1 STANDARD; PRT; 225 AA.
ID GP30_BPSP1
AC 038423;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Gene 30 protein (GP30).
GN 30.
OS Bacteriophage SP01.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Myoviridae;
OC SPO1-like viruses.
OX NCBI_TaxID=10685;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92267370; PubMed=1587473;
RA Scarlato V., Sayre M.H.;
RT "Sequence of the bacteriophage SP01 gene 30.";
RL Gene 114:115-119(1992).
CC -1- FUNCTION: ESSENTIAL FOR DNA SYNTHESIS.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL; M82842; AAA32597.1; -; 98442FCB05FAF7D CRC64;
SO SEQUENCE 225 AA; 25858 MW;

Query Match 87.1%; Score 27; DB 1; Length 225;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYED 5
Db 32 VYED 36

RESULT 17

YUBP_ECOLI

ID YUBP_ECOLI STANDARD; PRT; 273 AA.

AC P18005; P52149;

DT 01-NOV-1990 (Rel. 16, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical protein yubp.

GN YUBP OR YGBB.

OS Escherichia coli.

OC Plasmid F, and Plasmid IncFII R1.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.

OX NCBI_Taxid=562;

RN [1]

RP SEQUENCE FROM N.A.

RC PLASMID-F;

RX MEDLINE=90136505; PubMed=2693941;

RA Loh S., Gram D., Skurray R.;

RT "Nucleotide sequence of the leading region adjacent to the origin of
RT transfer on plasmid F and its conservation among conjugative
RT plasmids.";

RL Mol. Genet. 219:177-186(1989).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-K12 / CR63; PLASMID-F;

RA Shimizu H., Satoh Y., Suda Y., Uehara K., Sempel G., Mizobuchi K.;

RT "Complete nucleotide sequence of the F plasmid: its implications for
RT organization and diversification of plasmid genomes.";

RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

RL Nucleic Acids Res. 18:1046-1046(1990).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN-K12; PLASMID-IncFII R1;

RX MEDLINE=90192087; PubMed=2315017;

RA Graus H., Hoedel A., Wallner P., Hoegenauer G.;

RT "The sequence of the leading region of the resistance plasmid R1.";

FT VARIANT 210 210 W -> R (IN PLASMID INCFII R1).
FT CONFLICT 34 34 N -> T (IN REF. 1).
SO SEQUENCE 273 AA; 31615 MW; F136DC9182E80CFA CRC64;

DR EMBL; M82842; AAA32597.1; -; 98442FCB05FAF7D CRC64;
SO SEQUENCE 225 AA; 25858 MW;

Query Match 87.1%; Score 27; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYED 5
Db 189 VYED 193

RESULT 18

GDFP_HUMAN

ID GDFP_HUMAN STANDARD; PRT; 308 AA.

AC Q99968; P78360; O14629; Q9NRF0;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Growth/differentiation factor 15 precursor (GDF-15) (Placental bone
DE morphogenic protein) (Placental TGF-beta) (Macrophage inhibitory
DE cytokine-1) (MIF-1) (Prostate differentiation factor) (NSAID-regulated
DE protein 1) (NRG-1).

GN GDF15 OR PLAB OR PTGFB OR MIF1 OR PDF.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI_Taxid=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE-Fibrosarcoma;

RX MEDLINE=98006316; PubMed=9348093;

RA Yokoyama-Kobayashi M., Seeki M., Sekine S., Kato S.;

RT "Human cDNA encoding a novel TGF-beta superfamily protein highly
RT expressed in placenta.";

RL J. Biochem. 122:622-626(1997).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98041637; PubMed=9375789;

RA Thomas R., Huford M., Sutton J., Xu D., Li Y., Lu L.;

RT "PLAB, a novel placental bone morphogenetic protein.";

RL Biochim. Biophys. Acta 1354:40-44(1997).

RN [4]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98256302; PubMed=9593718;

RA Parakekar V.M., Vail A.L., Grasser W.A., Brown T.A., Xu H.;

RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";

RL J. Biol. Chem. 273:13760-13767(1998).

RN [5]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98085971; PubMed=9426002;

RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Qiu L., Baumes S.A.;

RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";

RL J. Biol. Chem. 273:13760-13767(1998).

RN [6]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98085971; PubMed=9426002;

RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Qiu L., Baumes S.A.;

RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";

RL J. Biol. Chem. 273:13760-13767(1998).

RN [7]

RP SEQUENCE FROM N.A.

RC TISSUE-Placenta;

RX MEDLINE=98085971; PubMed=9426002;

RA Lawton L.N., de Fatima Bonaldo M., Jelenc P.C., Qiu L., Baumes S.A.;

RT "Cloning and characterization of a novel member of the transforming
RT growth factor-beta/bone morphogenetic protein family.";

RA Baek S.J., Nixon J., Eling T.;
 RT "NRG-1 is associated with apoptosis."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- SUBMITTER: HOMODIMER; DISULFIDE-LINKED (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, WITH LOWER
 CC LEVELS IN PROSTATE AND COLON AND SOME EXPRESSION IN KIDNEY.
 CC -1- SIMILARITY: BELONGS TO THE TGF-BETA FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AB000584; BAA19151.1; -
 DR EMBL: AF019770; AAB88673.1; -
 DR EMBL: U88323; AAB88913.1; -
 DR EMBL: AF003934; AAC24456.1; -
 DR EMBL: AF008303; AAC39537.1; -
 DR EMBL: AF173860; AAF89834.1; -
 DR HSSP: P18075; IAMP.
 DR MIM: 605312; -
 DR InterPro: IPR002400; GFCysknot.
 DR InterPro: IPR001839; TGFb.
 DR Pfam: PF00019; TGF-Beta; 1.
 DR PRINTS: PR00438; GFCYSKNOT.
 DR ProDom: PD000357; TGFb; 1.
 DR SMART: SM00204; TGFb; 1.
 DR PROSITE: PS00250; TGF-BETA; 1; FALSE NEG.
 KW Growth factor; Cytokine; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 29
 FT PROPEP 30 194
 FT CHAIN 195 308
 FT DISULFID 211 274
 FT DISULFID 240 305
 FT DISULFID 244 307
 FT DISULFID 273 273
 FT CARBOHD 70 70
 FT VARIANT 48 48
 FT FTID=VAR_010386.
 FT CONFLICT 9 9
 FT CONFLICT 202 202
 FT CONFLICT 269 269
 FT CONFLICT 288 288
 FT SEQUENCE 308 AA; 34168 MW; ADF3A3ED065ACAZE CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 308;
 Best Local Similarity 100.0%; Pred. No. 68;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 RYEDL 6
 Db 59 RYEDL 63
 RESULT 19
 MRAM_BACSU
 ID MRAM_BACSU STANDARD; PRT; 311 AA.
 AC 007876;
 DT 01-OCT-1994 (Rel. 30, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE S-adenosyl-methyltransferase mram (EC 2.1.1.-).
 GN MRAM.
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;

RA Daniel R.A., Williams A.M., Errington J.;
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE OF 108-311 FROM N.A.
 RC STRAIN=168;
 RA MEDLINE=94064553; PubMed=8244929;
 RX Yanour A., Daniel R.A., Errington J., Buchanan C.E.;
 RT "Cloning and sequencing of the cell division gene ppgB, which encodes
 RT penicillin-binding protein 2B in Bacillus subtilis."
 RL J. Bacteriol. 175:7604-7616(1993).
 CC -1- FUNCTION: Exhibits a S-adenosyl-dependent methyltransferase
 CC activity (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE MRAM FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: Z68230; CAA92525.1; -
 DR EMBL: L09703; AAC36835.1; -
 DR EMBL: Z59111; CAB13587.1; -
 DR Subtilist; BG10219; mram.
 DR InterPro: IPR002903; Bac_Metnfrase.
 DR Pfam: PF01795; Methyltransf.5; 1.
 DR ProDom: PD004685; Bac_Metnfrase; 1.
 DR TIGRPFAMS: TIGR00006; UPE0117; 1
 KW Transferase; Methyltransferase; Complete proteome.
 SEQUENCE 311 AA; 35305 MW; C50C609129DFDAD CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 311;
 Best Local Similarity 100.0%; Pred. No. 89;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 2 RYEDL 6
 Db 143 RYEDL 147
 RESULT 20
 REBB_NEIMA
 ID REBB_NEIMA STANDARD; PRT; 341 AA.
 AC Q9S642;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE drdp-glucose 4,6-dehydratase (EC 4.2.1.46).
 GN (REBB) OR NMA0189) AND (REBB2 OR NMA0204).
 OS Neisseria meningitidis (serogroup A).
 OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
 OX NCBI_TaxID=65699;
 RN [1]
 RP SEQUENCE FROM N.A.
 RP STRAIN=MA-1 / Serogroup A;
 RC MEDLINE=99150278; PubMed=10024588;
 RA Lee F.K., Gibson B.W., Melanagh W., Zaleski A., Apicella M.A.;
 RT "Relationship between UDP-glucose 4-epimerase activity and
 RT oligoglucose glycoforms in two strains of Neisseria meningitidis."
 RL Infect. Immun. 67:1405-1414(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP STRAIN=Z2491 / Serogroup A / Serotype 4A;
 RC MEDLINE=20222556; PubMed=10761919;
 RA Parkhill J., Achtman M., James K.D., Bentley S.D., Churcher C.,
 RA Klee S.R., Morelli G., Basham D., Brown D., Chillingworth T.,
 RA Davies R.M., Davis P., Devlin K., Feltwell T., Hamlin N., Holroyd S.,
 RA Jagers K., Leather S., Moule S., Mungall K., Quail M.A.,
 RA Rajandream M.A., Rutherford K.M., Symonds M., Skelton J.,
 RA Whitehead S., Spratt B.G., Barrett B.G.;
 RT "Complete DNA sequence of a serogroup A strain of Neisseria

RT meningitidis 22491.1;
RL Nature 404:502-506(2000).
CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
CC glucose + H(2)O.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: DTPD-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTPD-GLUCOSE
CC DEHYDRATASE SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AF083467; AAD23919.1; -
CC EMBL: AL162752; CAB83504.1; -
CC EMBL: AL162752; CAB83518.1; -
CC HSSP: P27830; IBXK.
CC InterPro: IPR001509; Epimerase_Dh.
CC Pfam: PF01370; Epimerase; 1.
CC TIGRfams: TIGR01181; dTDP-gluc_dehyd; 1.
CC Lipopolysaccharide biosynthesis; Lyase; NAD; Complete proteome.
CC NP_BIND 8 NAD (POTENTIAL).
CC SEQUENCE 341 AA; 38224 MW; 826023B63916E72A CRC64;
SQ
Query Match 87.1%; Score 27; DB 1; Length 341;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 RYEDL 6
Db 283 RYEDL 287
RESULT 21
REPB_NEIGO STANDARD; PRT; 346 AA.
ID REPB_NEIGO
RT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE 30-MAY-2000 (Rel. 39, Last annotation update)
DE dTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
GN REPB.
OS Neisseria gonorrhoeae.
OC Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
CX NCBI_TaxID=485;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=MS11;
RX MEDLINE=95050260; PubMed=7961452;
RA Robertson B.D., Frosch M., van Putten J.P.M.;
RA "Complete identification of cryptic rhamnose biosynthesis genes in
RA Neisseria gonorrhoeae and their relationship to lipopolysaccharide
RT biosynthesis.";
RL J. Bacteriol. 176:6915-6920(1994).
CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
CC glucose + H(2)O.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: DTPD-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTPD-GLUCOSE
CC DEHYDRATASE SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).
CC
CC EMBL: 232742; CAA83652.1; -
CC EMBL: 221508; CAA79718.1; -
CC PIR: S47045; S47045.
CC HSSP: P27830; IBXK.
CC InterPro: IPR001509; Epimerase_Dh.
CC Pfam: PF01370; Epimerase; 1.
CC TIGRfams: TIGR01181; dTDP-gluc_dehyd; 1.
CC Lipopolysaccharide biosynthesis; Lyase; NAD.
CC NP_BIND 13 NAD (POTENTIAL).
CC SEQUENCE 346 AA; 38828 MW; B52FD8125C567675 CRC64;
SQ
Query Match 87.1%; Score 27; DB 1; Length 346;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 2 RYEDL 6
Db 288 RYEDL 292
RESULT 22
REPB_NEIMB STANDARD; PRT; 355 AA.
ID REPB_NEIMB
RT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 40, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE dTDP-glucose 4,6-dehydratase (EC 4.2.1.46).
DE (REPB1 OR NMB0063) AND (REPB2 OR NMB0079).
GN Neisseria meningitidis (serogroup B).
OS Bacteria; Proteobacteria; beta subdivision; Neisseriaceae; Neisseria.
CX NCBI_TaxID=491;
RN
RN SEQUENCE FROM N.A.
RC STRAIN=B1940 / Serogroup B;
RX MEDLINE=944293762; PubMed=8022285;
RA Hammerschmidt S., Birkholz C., Zahrlinger U., Robertson B.D.,
RA van Putten J.P.M., Ebeling O., Frosch M.;
RA "Contribution of genes from the capsule gene complex (cps) to
RT lipopolysaccharide biosynthesis and serum resistance in Neisseria
RT meningitidis.";
RL Mol. Microbiol. 11:885-896(1994).
RN
RN SEQUENCE FROM N.A.
RC STRAIN=MC58 / Serogroup B;
RX MEDLINE=20175755; PubMed=10710307;
RA Tettelin H., Saunders N.J., Heidelberg J., Jeffries A.C., Nelson K.E.,
RA Eisen J.A., Ketchum K.A., Hood D.W., Peden J.F., Dodson R.J.,
RA Nelson W.C., Gwin M.L., Deboy R., Peterson J.D., Hickey E.R.,
RA Haft D.H., Salzberg S.L., White O., Fleischmann R.D., Dougherty B.A.,
RA Mason T., Ciecko A., Parksey D.S., Blair E., Clifton H., Clark E.B.,
RA Cotton M.D., Utterback T.R., Khouri H., Qin H., Vamathevan J.,
RA Gill J., Scarlata V., Maignani V., Pizzari G., Grandi G., Sun L.,
RA Smith H.O., Fraser C.M., Moxon E.R., Rappuoli R., Venter J.C.;
RA "Complete genome sequence of Neisseria meningitidis serogroup B strain
RT MC58.";
RL Science 287:1809-1815(2000).
CC -1- CATALYTIC ACTIVITY: dTDP-glucose = dTDP-4-dehydro-6-deoxy-D-
CC glucose + H(2)O.
CC -1- COFACTOR: NAD.
CC -1- PATHWAY: DTPD-L-RHAMNOSE BIOSYNTHESIS WITHIN THE O ANTIGEN
CC BIOSYNTHESIS PATHWAY OF LIPOPOLYSACCHARIDE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE SUGAR EPIMERASE FAMILY. DTPD-GLUCOSE
CC DEHYDRATASE SUBFAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

```

CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: I09188; AAA63157.1; -.
DR EMBL: AE002365; AAP40531.1; -.
DR EMBL: AE002367; AAP40543.1; -.
DR HSSP: P27830; IBXK.
DR TIGR: NMB0063; -.
DR TIGR: NMB0079; -.
DR InterPro: IPR001509; Epimerase_Dh.
DR Pfam: PF01370; Epimerase; 1.
DR TIGRfam: TIGR01181; drdp_gluc_dehyd; 1.
KM Lipopolysaccharide biosynthesis; Lyase; NAD; Complete proteome.
FT NP_IND 8 14 NAD (POTENTIAL).
FT CONFLICT 1 3 MKK -> MOTANKRT (IN REF. 1).
FT CONFLICT 26 28 RDA -> ODS (IN REF. 1).
FT CONFLICT 32 32 V -> L (IN REF. 1).
FT CONFLICT 46 47 EV -> DI (IN REF. 1).
FT CONFLICT 73 73 Y -> H (IN REF. 1).
FT CONFLICT 142 142 G -> H (IN REF. 1).
FT CONFLICT 152 152 A -> T (IN REF. 1).
FT CONFLICT 268 268 A -> T (IN REF. 1).
FT CONFLICT 274 274 A -> V (IN REF. 1).
SQ SEQUENCE 355 AA; 39865 MW; 631AAOEDPA02B6F41 CRC64;

Query Match
Best Local Similarity 87.1%; Score 27; DB 1; Length 355;
Matches 5; Conservative 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 283 RYEDL 287

RESULT 23
PKG_STAM STANDARD: PRI: 396 AA.
AC Q9ZSC4;
DI 30-MAY-2000 (Rel. 39, Created)
DI 30-MAY-2000 (Rel. 39, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphoglycerate kinase (EC 2.7.2.3).
GN PGK OR SAV0773 OR SA0728 OR MM0735.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315),
OS Staphylococcus aureus (strain MW2), and
OS Staphylococcus aureus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158876, 158879, 196620, 1280;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Iian J.-O., Ito T.,
RA Kanamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyama A.,
RA Mizutani-Oi Y., Takahashi N.K., Sawano T., Inoue R.-I., Kato C.,
RA Sekimizu K., Hirakawa H., Kuhara S., Goto S., Yabuzaki J.,
RA Kanehisa M., Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RA "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus.";
RT Lancet 357:1225-1240(2001).
RL [2]
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-MW2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takeuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naito T., Kuroda H., Cui L.,
RA Yamamoto K., Hiramatsu K.;
RA "Genome and virulence determinants of high virulence community-
RT acquired MRSA.";
RT Lancet 359:1819-1827(2002).
RL [3]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN-BB;
RA Morrissey J.A., Williams P.;
RT "Isolation and characterisation of a glycolytic operon in
RT Staphylococcus aureus.";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -1 CATALYTIC ACTIVITY: ATP + 3-phospho-D-glycerate = ADP + 3-
CC phospho-D-glyceroyl phosphate.
CC -1 PATHWAY: Second phase of glycolysis; second step.
CC -1 SUBUNIT: MONOMER (BY SIMILARITY).
CC -1 SUBCELLULAR LOCATION: Cytoplasmic.
CC -1 SIMILARITY: BELONGS TO THE PHOSPHOGLYCERATE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@lsb.sib.ch).
CC -----
DR EMBL: AP003360; BAB56935.1; -.
DR EMBL: AP003361; BAB41961.1; -.
DR EMBL: AP004824; BAB4600.1; -.
DR EMBL: AJ133520; CAB38646.1; -.
DR HSSP: P36204; IYPE.
DR InterPro: IPR001576; PGK.
DR Pfam: PF00162; PGK; 1.
DR PRINTS: PR00477; PHGLYKINASE.
DR PROSITE: PS00111; PHGLYCERATE_KINASE; 1.
KM Transferrase; Kinase; Glycolysis; Complete proteome.
SQ SEQUENCE 396 AA; 42601 MW; C8383CFBDCB5F4EC CRC64;

Query Match
Best Local Similarity 87.1%; Score 27; DB 1; Length 396;
Matches 5; Conservative 100.0%; Pred. No. 1.2e+02; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 24
LE21_ARCFU STANDARD: PRI: 418 AA.
ID LE21_ARCFU
AC O28316;
DI 15-JUL-1998 (Rel. 36, Created)
DI 15-JUN-2002 (Rel. 41, Last sequence update)
DI 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-Isopropylmalate dehydratase large subunit 1 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 1) (Alpha-IRM isomerase 1) (IPMI 1).
GN LE21 OR AF1963.
OS Archaeoglobus fulgidus.
OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
OC Archaeoglobaceae; Archaeoglobus.
OX NCBI_TaxID=2234;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-VC-16 / DSM 4304 / ATCC 49558;
RX MEDLINE=98045343; PubMed=9389475;
RA Kleink H.-P., Clayton R.A., Tomb J.-F., White O., Nelson K.E.,
RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.K., Peterson J.D.,
RA Richardson D.L., Kerlavage A.R., Graham D.E., Kyrpides N.C.,
RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
RA Kirkness E.F., Dougherty B.A., McKenney K., Adams M.D., Loftus B.,
RA Peterson S., Ketch C.L., McNeil L.K., Badger J.H., Glodek A., Zhou L.,
RA Overbeek R., Gocayne J.D., Weidman J.F., McDonald L., Utterback T.,
RA Cotton M.D., Spriggs T., Artiach P., Kaine B.P., Sykes S.M.,
RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
RA Venter J.C.;
RT "The complete genome sequence of the hyperthermophilic, sulphate-
RT reducing archaeon Archaeoglobus fulgidus.";

```

RL Nature 390:364-370(1997).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE000967; AAB89290.1; ALT_INT.
DR TIGR: AF1963.
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF00330; aconitase_1.
DR ProDom: PD000511; Aconitase_N; 1.
DR PROSITE: PS00450; ACONITASE_1; 1.
DR PROSITE: PS01244; ACONITASE_2; 1.
KM Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 298 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 358 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 361 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 418 AA; 45140 MW; 51B2BD0A3D9C9A22 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 302 RYEDL 306
|||||
LE21_METKA STANDARD; PRT; 418 AA.
ID LE21_METKA
AC O8TVF2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydratase large subunit 1 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 1) (Alpha-IPM isomerase 1) (IPMI 1).
GN LEUC1 OR MKJ440.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
CC Methanopyrus.
CC NCBI_TaxID=2320;
CX NCBI_TaxID=2320;
RX MEDLINE=21927647; Pubmed=11930014;
RA Slesarev A.I., Mezheva K.V., Makarova K.S., Polushin N.N.,
RA Shcherbina O.V., Shakhova V.V., Belova G.I., Atavudin L.,
RA Natale D.A., Rozozin I.B., Telusov R.L., Wolf Y.I., Stettin K.O.,
RA Malykh A.G., Koonin E.V., Kozayavkin S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.

CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE010436; AM02653.1;
DR InterPro: IPR001030; Aconitase_N.
DR Pfam: PF00330; aconitase_1.
DR ProDom: PD000511; Aconitase_N; 1.
DR PROSITE: PS00450; ACONITASE_1; 1.
DR PROSITE: PS01244; ACONITASE_2; FALSE_NEG.
KM Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
FT METAL 298 298 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 358 358 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 361 361 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 418 AA; 45202 MW; FABA735A9BD51C72 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 418;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 302 RYEDL 306
|||||
LE22_METTH STANDARD; PRT; 428 AA.
ID LE22_METTH
AC O27668;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydratase large subunit 2 (EC 4.2.1.33)
DE (Isopropylmalate isomerase 2) (Alpha-IPM isomerase 2) (IPMI 2).
GN LEUC2 OR MTH1631.
OS Methanobacterium thermoautotrophicum.
OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;
CC Methanobacteriaceae; Methanothermobacter.
CC NCBI_TaxID=187420;
CX NCBI_TaxID=187420;
RX MEDLINE=98037514; Pubmed=9371463;
RA Smith D.R., Doucet-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,
RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,
RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,
RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,
RA Jivani N., Caruso A., Bush D., Safer H., Patwell D., Pradhanak S.,
RA Danaleis C.J., Mac J.-I., Rice P., Noelling J., Reeve G.M.;
RT "Complete genome sequence of Methanobacterium thermoautotrophicum
RT deltaH: functional analysis and comparative genomics.";
RL J. Bacteriol. 179:7135-7155(1997).
CC -1- FUNCTION: Catalyzes the isomerization between 2-isopropylmalate
CC and 3-isopropylmalate, via the formation of 2-isopropylmalate.
CC -1- CATALYTIC ACTIVITY: 3-isopropylmalate = 2-isopropylmalate +
CC H₂O.
CC -1- CATALYTIC ACTIVITY: 2-isopropylmalate + H₂O = 2-
CC isopropylmalate.
CC -1- PATHWAY: Leucine biosynthesis; second step.
CC -1- SUBUNIT: Heterodimer of leuc and leud (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ACONITASE/IPM ISOMERASE FAMILY. LEUC 2
CC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE000922; AAB6104.1; -
 DR InterPro: IPR001030; Aconitase_N.
 DR Pfam: PF000330; aconitase_1.
 DR ProDom: PD000511; Aconitase_N; 1.
 DR PROSITE; PS00450; ACONITASE_1; 1.
 DR PROSITE; PS01244; ACONITASE_2; 1.
 KW Leucine biosynthesis; Lyase; Iron-sulfur; 4Fe-4S; Complete proteome.
 FT METAL 304 304 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 364 364 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 FT METAL 367 367 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
 SQ SEQUENCE 428 AA; 46447 MW; B93FIDC53C9B8178 CRC64.

Query Match 87.1%; Score 27; DB 1; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 308 RYEDL 312

RESULT 27
 DCDA_BACHD STANDARD; PRT: 439 AA.
 AC O9KCM5;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Diaminopimelate decarboxylase (EC 4.1.1.20) (DAP decarboxylase).
 GN LYSA OR BH1544.
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_Taxid:86665;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C-125 / JCM 9153;
 RX MEDLINE=20512582; PubMed=11058132;
 RA Takami H., Nakasone K., Takaki Y., Mieno G., Sasaki R., Masui N.,
 RA Fujii F., Hirama C., Nakamura Y., Ogatawara N., Kihara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).
 CC -1- CATALYTIC ACTIVITY: Meso-2,6-diaminoheptanedioate = L-lysine +
 CC CO(2).
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: Lysine biosynthesis; last step.
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF ORNITHINE, DAP, AND ARGININE
 CC DECARBOXYLASES.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AP001512; BAB05263.1; -
 DR InterPro: IPR000183; Decarboxylase2.
 DR Pfam; PF00278; Orn_DAP_Arg_dec; 1.
 DR Pfam; PF02784; Orn_Arg_dec_N; 1.
 DR PRINTS; PR01179; ODADCRXLASE.
 DR TIGRFAMS; TIGR01048; lysA; 1.
 DR PROSITE; PS00878; ODR_DC_2.1; FALSE_NEG.

DR PROSITE; PS00879; ODR_DC_2.2; 1.
 KW Lysine biosynthesis; Lyase; Decarboxylase; Pyridoxal phosphate;
 KW Complete proteome.
 FT BINDING 66 66 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 SQ SEQUENCE 439 AA; 48444 MW; FE59E2103BEE90F CRC64;

Query Match 87.1%; Score 27; DB 1; Length 439;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 422 RYEDL 426

RESULT 28
 GUAD_ECOLI STANDARD; PRT: 439 AA.
 ID GUAD_ECOLI
 AC P76641; Q46816;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Guanine deaminase (EC 3.5.4.3) (Guanase) (Guanine aminase) (Guanine
 DE aminohydrolase) (GAH).
 GN GUAD OR B2883.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Escherichia.
 OX NCBI_Taxid:562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K12 / MG1655.
 RX MEDLINE=9742617; PubMed=9278503;
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 RA Mau B., Shao Y.;
 RT "The complete genome sequence of Escherichia coli K-12.";
 RL Science 277:1453-1474(1997).
 RN [2]
 RP CHARACTERIZATION.
 RX MEDLINE=20372655; PubMed=10913105;
 RA Maynes J.T., Yuan R.G., Snyder F.F.;
 RT "Identification, expression, and characterization of Escherichia coli
 RT guanine deaminase.";
 RL J. Bacteriol. 182:4658-4660(2000).
 CC -1- FUNCTION: CATALYZES THE HYDROLYTIC DEAMINATION OF GUANINE,
 CC PRODUCING XANTHINE AND AMMONIA.
 CC -1- CATALYTIC ACTIVITY: Guanine + H(2)O = xanthine + NH(3).
 CC -1- COFACTOR: CONTAINS 1 MOLE OF ZINC PER SUBUNIT.
 CC -1- PATHWAY: PURINE NUCLEOTIDE CATABOLISM.
 CC -1- SIMILARITY: BELONGS TO THE ATZ/TRZ FAMILY.

CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC -----
 DR EMBL; AE000372; AAC75921.1; -
 DR EMBL; U28375; AAA83064.1; -
 DR Ecocore; EGI3066; guad.
 DR InterPro: IPR002604; ATZ_TRZ.
 DR Pfam; PF01685; ATZ_TRZ; 1.
 KW Hydrolyase; Zinc; Complete proteome.
 SQ SEQUENCE 439 AA; 50244 MW; 43389F3AF9E4AD83 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 439;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
 DB 112 RYEDL 116

RESULT 29
 C133_DROME STANDARD; PRT; 492 AA.
 ID C133_DROME
 AC 09VGB3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Probable cytochrome P450 313a3 (EC 1.14.-.-) (CYPCXCXIII1A3).
 GN CYP313A3 OR CG10093.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachiycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 NX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mikos G.L.G.,
 RA Abail J.F., Agbayani A., An H.-J., Andrews-Plannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
 RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabriellian A.E., Garb N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris J.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howard T.J., Wei M.-H., Idegum C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 RN [2]
 RP CONCEPTUAL TRANSLATION.
 RA Nelson B.;
 RL Unpublished observations (SEP-2000).
 CC -1- FUNCTION: MAY BE INVOLVED IN THE METABOLISM OF INSECT HORMONES AND
 CC IN THE BREAKDOWN OF SYNTHETIC INSECTICIDES (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) -> ROH +
 CC oxidized flavoprotein + H(2)O.
 CC -1- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum
 CC (potential).
 CC -1- SIMILARITY: BELONGS TO THE CYTOCHROME P450 FAMILY.
 CC -1- CAUTION: Ref.1 sequence differs from that shown due to erroneous

CC gene model prediction.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as the content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.slb.ch/announce/>
 CC or send an email to license@isb.slb.ch).
 CC -----
 CC EMBL: AE003695; AAF54770.2; ALT_SEQ.
 CC FLYBASE: Fg00038007; CYP313a3.
 CC InterPro: IPR001128; Cytochrome_P450.
 CC Pfam: PF00067; P450.1.
 CC PRINTS: PR00385; P450.
 CC PROSITE: PS00086; CYTOCHROME_P450.1.
 CC OXIDOREDUCTASE; Monooxygenase; Membrane; Heme; Microsome;
 CC Endoplasmic reticulum; Hypothetical protein.
 CC BINDING 438 HEME (BY SIMILARITY).
 CC SEQUENCE 492 AA; 56195 MW; 2836DJ3953CB4DC CRC64;
 SO

Query Match 87.1%; Score 27; DB 1; Length 492;
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
 DB 466 RYEDL 470

RESULT 30
 FUT3_ARATH
 ID FUT3_ARATH STANDARD; PRT; 493 AA.
 AC 09CA71;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Probable fucosyltransferase 3 (EC 2.4.1.-) (ATFUT3).
 GN FUT3 OR ATIG74420 OR FIM20.10.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 NX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608393; PubMed=11743104;
 RA Sarría R., Wagner T.A., O'Neill M.A., Falk A., Wilkerson C.G.,
 RA Keegstra K., Raikhel N.Y.;
 RT Characterization of a family of Arabidopsis genes related to
 RL xyloglucan fucosyltransferase1."
 RL Plant Physiol. 127:1595-1606(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Columbia;
 RX MEDLINE=21016719; PubMed=11130712;
 RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
 RA White O., Alonso J., Altafi H., Araujo R., Bowman C.L., Brooks S.Y.,
 RA Buehler E., Chan A., Chao Q., Chen H., Cheuk R.F., Chin C.W.,
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
 RA Dunn P., Egu P., Feldhahn T.V., Feng J.-D., Fong B., Fujii C.Y.,
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huizlar L.,
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
 RA Kim C.U., Koo H.L., Kremenetskaia I., Kutz D.B., Kwan A., Lam B.,
 RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
 RA Lin X., Liu S.X., Liu Z.A., Luros J.S., Malit R., Marshall A.,
 RA Miltner J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
 RA Pai G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
 RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Souhwaik A.M.,
 RA Sun H., Talbot L.J., Tambunga G., Tatum M.J., Town C.D.,
 RA Utterback T., Van Aken S., Vaysberg M., Vysotskaia V.S., Walker M.,
 RA Wu D., Yu G., Fraser C.M., Vaynsberg J.C., Davis R.W.;
 RT "Sequence and analysis of chromosome 1 of the plant *Arabidopsis*

```

RT thaliana."
RL Nature 408:816-820(2000).
CC -1- FUNCTION: May be involved in cell wall biosynthesis. May
CC act as a fucosyltransferase.
CC -1- PATHWAY: Glycosylation.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Membrane-bound
CC form in trans cisternae of Golgi (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in roots, stems, leaves, flowers,
CC siliques and seedlings.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 37.
CC -1- CAUTION: THE SEQUENCE IN REF.2 DIFFERS FROM THAT SHOWN DOE TO
CC WRONG EXON PREDICTIONS FROM THE GENOMIC SEQUENCE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF417473; AAL50622.1; -.
DR EMBL: AC011765; AAG52352.1; ALU_SEQ.
DR InterPro: IPR004938; XG_Ftase.
DR Pfam: PF03254; XG_Ftase; 1.
KW Transferase; glycosyltransferase; Transmembrane; glycoprotein;
KW Signal-anchor; Golgi stack; Cell wall.
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 12 32 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT FT (POTENTIAL).
FT DOMAIN 33 493 MEMBRANAL, CATALYTIC (POTENTIAL).
FT POLY-SER.
FT CARBOHYD 124 124 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 493 AA; 55997 MW; 64CB98C78C26DEA3 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 493;
Best Local Similarity 100.0%; Pred. NO. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 108 RYEDL 112

RESULT 31
Y213_HAEIN
ID Y213_HAEIN STANDARD: PRT; 514 AA.
AC P44572:
DT 01-NOV-1995 (rel. 32, Created)
DT 01-NOV-1995 (rel. 32, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Putative binding protein H10213 precursor.
GN H10213.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OC NCBI_TaxID=727;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KM20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kerlavage A.R., Bult C.J., Tomb J.F., Dougherty B.A., Merrick J.M.,
RA Scott K., Sutton G., Fiteugh W., Fields C.A., Gocayne J.D.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Utechtack T.R., Hanna M.C., Nguyen D.T., Saudek D.N., Brandon R.C.,
RA Fine L.D., Fritchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae

```

```

RT Rd";
RL Science 269:496-512(1995).
CC -1- FUNCTION: PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
CC (Probable).
CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
CC PROTEIN FAMILY 5.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32706; AAC21881.1; -.
DR HSSP: P06202; 1B52.
DR TIGR: H10213; -.
DR InterPro: IPR000437; Prok_Lipoprot.
DR InterPro: IPR000914; SBP_bac_5.
DR Pfam: PF00496; SBP_bac_5; 1.
DR PROSITE: PS01040; SSB_BACTERIAL_5; 1.
DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; FALSE_NEG.
KW Hypothetical protein; Transport; Membrane; Signal; Lipoprotein;
KW Complete proteome.
FT SIGNAL 1 23 POTENTIAL.
FT CHAIN 24 514 PUTATIVE BINDING PROTEIN H10213.
FT LIPID 24 24 N-ACYL DIGLYCERIDE (PROBABLE).
SQ SEQUENCE 514 AA; 58876 MW; 789188C4358BDEBC CRC64;

Query Match 87.1%; Score 27; DB 1; Length 514;
Best Local Similarity 66.7%; Pred. NO. 1.5e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 359 IRYDDL 364

RESULT 32
PPCK_PYRO
ID PPCK_PYRO STANDARD: PRT; 621 AA.
AC O58050;
DT 16-OCT-2001 (rel. 40, Created)
DT 16-OCT-2001 (rel. 40, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32) (PEP
DE carboxylase) (Phosphoenolpyruvate carboxylase) (PEPCK).
GN PCK OR PHO312.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OC NCBI_TaxID=53953;
RN NCB1
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RX MEDLINE=96344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosooya A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamaya M., Ohiku Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Kosb F.T., Horikoshi K.,
RA Masuchi Y., Shizuya H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
RT thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";
RL DNA Res. 5:55-76(1998).
CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate
CC + CO(2);
CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity).
CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]

```

CC FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AP000001; BAA29385.1; -

CC InterPro: IPR000364; PEP_carboxykin.

CC Pfam: PF00821; PEPCK.1.

CC ProDom: PD004738; PEP_carboxykin.1.

CC PROSITE: PS00505; PEPCK.GTP; FALSE.NEG.

CC Glucocoenogenesis; Lyase; Decarboxylase; GTP-binding; Complete proteome.

CC NP_BIND 218 225 GTP (POTENTIAL).

CC ACT_SITE 269 269 BY SIMILARITY.

CC SEQUENCE 621 AA; 72201 MW; 738232FF30F20F01 CRC64;

CC

CC Query Match 87.1%; Score 27; DB 1; Length 621;

CC Best Local Similarity 100.0%; Pred. No. 1.9e+02;

CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

CC 2 RYEDL 6

CC Db 537 RYEDL 541

CC

CC RESULT 33

CC SKNI_CABEL STANDARD; PRT; 623 AA.

CC AC P34707;

CC DT 01-FEB-1994 (Rel. 28, Created)

CC DT 15-JUN-2002 (Rel. 41, Last sequence update)

CC DT 15-JUN-2002 (Rel. 41, Last annotation update)

CC DE Skn-1 protein.

CC SKN-1 OR T19E7.2.

CC Caenorhabditis elegans.

CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

CC OC Rhabditidae; Peloderae; Caenorhabditis.

CC OX NCBI_TaxID=6239;

CC [1]

CC RP SEQUENCE FROM N.A.

CC RC STRAIN-Bristol N2;

CC RA Pauley A., Galtung S.;

CC RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.

CC RN [2]

CC RP REVISIONS.

CC RA Waterston R.;

CC RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.

CC RN [3]

CC RP SEQUENCE OF 49-623 FROM N.A., FUNCTION, AND DEVELOPMENTAL STAGE.

CC RC STRAIN-Bristol N2;

CC RX MEDLINE=92191285; PubMed=1547503;

CC RA Bowerman B., Eaton B.A., Pless J.R.;

CC RT "Skn-1, a maternally expressed gene required to specify the fate of

CC ventral blastomeres in the early C. elegans embryo.";

CC RL Cell 68:1061-1075(1992).

CC RN [4]

CC RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 546-619.

CC RX MEDLINE=98290451; PubMed=9628487;

CC RA Rupert P.B., Daughdrill G.W., Bowerman B., Matthews B.W.;

CC RT "A new DNA-binding motif in the Skn-1 binding domain-DNA complex.";

CC RL Nat. Struct. Biol. 5:484-491(1998).

CC CC -1- FUNCTION: REQUIRED TO SPECIFY THE FATE OF VENTRAL BLASTOMERES IN

CC THE EARLY EMBRYO, AND POSTEMBRYONICALLY FOR THE DEVELOPMENT OF THE

CC INTERSTINE.

CC CC -1- SUBCELLULAR LOCATION: Nuclear (Probable).

CC CC -1- DEVELOPMENTAL STAGE: EARLY EMBRYOGENESIS, AND POSTEMBRYONIC

CC CC INTERSTINAL CELLS. EXPRESSED MATERNALLY.

CC CC -1- SIMILARITY: BELONGS TO THE SKNI FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: U42843; AAA83594.2; -

CC EMBL: M64359; -; NOT_ANNOTATED.CDS.

CC DR PIR; A42143; A42143.

CC DR PDB; 1SRN; 17-JUN-98.

CC DR TRANSFAC; T01614; -

CC DR WormRep; T19E7.2; CE27591.

CC DR InterPro: IPR004827; TF_bZIP.

CC DR PROSITE: PS00036; BZIP_BASIC.1.

CC KW Developmental protein; DNA-binding; Nuclear protein; 3D-structure.

CC FT DNA_BIND 597 612 BASIC MOTIF.

CC SEQUENCE 623 AA; 70708 MW; 31A70303AB7CC691 CRC64;

CC

CC Query Match 87.1%; Score 27; DB 1; Length 623;

CC Best Local Similarity 100.0%; Pred. No. 1.9e+02;

CC Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

CC 2 RYEDL 6

CC Db 240 RYEDL 244

CC

CC RESULT 34

CC PCK_PYRFU STANDARD; PRT; 624 AA.

CC ID PCK_PYRFU

CC AC Q80410;

CC DT 15-JUN-2002 (Rel. 41, Created)

CC DT 15-JUN-2002 (Rel. 41, Last sequence update)

CC DT 15-JUN-2002 (Rel. 41, Last annotation update)

CC DE Phosphoenolpyruvate carboxylase [GTP] (EC 4.1.1.32) (PEP

CC carboxylase) (Phosphoenolpyruvate carboxylase) (PEPCK).

CC PCK OR P70283.

CC GN Pyrococcus furiosus.

CC OS Pyrococcus furiosus.

CC OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

CC OC Pyrococcus.

CC OX NCBI_TaxID=2261;

CC [1]

CC RP SEQUENCE FROM N.A.

CC RC STRAIN-Vcl / DSM 3638 / ATCC 43587 / JCM 8422;

CC RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;

CC RT "The complete sequence of the Pyrococcus furiosus genome.";

CC RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

CC CC -1- CATALYTIC ACTIVITY: GTP + oxaloacetate = GDP + phosphoenolpyruvate

CC + CO(2).

CC CC -1- PATHWAY: Rate-limiting gluconeogenic enzyme.

CC CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC CC -1- SIMILARITY: BELONGS TO THE PHOSPHOENOLPYRUVATE CARBOXYKINASE [GTP]

CC FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AE010153; AAL80413.1; -

CC DR PROSITE: PS00505; PEPCK.GTP; FALSE.NEG.

CC KW Glucocoenogenesis; Lyase; Decarboxylase; GTP-binding; Complete proteome.

CC FT NP_BIND 223 230 GTP (POTENTIAL).

CC FT ACT_SITE 274 274 BY SIMILARITY.

CC SEQUENCE 624 AA; 72665 MW; EC6D925CF57EAC25 CRC64;

CC

CC Query Match 87.1%; Score 27; DB 1; Length 624;

CC Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 542 RYEDL 546

RESULT 35

GIDA_BUCAI STANDARD: PRT: 628 AA.

AC P57117; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glucose inhibited division protein A.
GN GIDA OR BU001.
OS Buchnera aphidicola (subsp. Acyrthosiphon pisum) (Acyrthosiphon pisum symbiotic bacterium).
OC Bacteria; Proteobacteria; gamma subdivision; Buchnera.
OX NCBI_TaxID=118099;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Tokyo 1998;
RX MEDLINE=20445173; PubMed=10993077;
RA Shigenobu S., Watanabe H., Hattori M., Sakaki Y., Ishikawa H.;
RT "Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS."
RL Nature 407:81-86(2000).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SIMILARITY: BELONGS TO THE GIDA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AP001118; BAB12729.1; -
DR InterPro: IPR002218; GIDA.
DR InterPro: IPR004416; GIDA_sub.
DR Pfam: PF01134; GIDA; 1.
DR Pfam: PF003738; GIDA; 1.
DR TIGRfam: TIGR00136; gida; 1.
DR PROSITE: PS01280; GIDA_1; 1.
DR PROSITE: PS01281; GIDA_2; 1.
KW Complete proteome.
SQ SEQUENCE 628 AA; 70307 MW; 0653745332BC077B CMC64;

Query Match 87.1%; Score 27; DB 1; Length 628;
Best Local Similarity 66.7%; Pred. No. 1.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RYEDL 6
DB 518 RYEDL 523

PEEL_HUMAN STANDARD: PRT: 653 AA.

AC 014829; 015253; 09JUT0; 09JUT2;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Serine/threonine protein phosphatase with EF-hands-1 (EC 3.1.3.16) (PPEF-1) (Protein phosphatase with EF calcium-binding domain) (PPEF) (Serine/threonine protein phosphatase 7) (PP7).
GN PPEF1 OR PPEF OR PP7C.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]
RP SEQUENCE FROM N.A.
RC Tissue-Retina:
RX MEDLINE=97471020; PubMed=9326663;
RA Sherman P.M., Sun H., Macke J.P., Williams J., Smallwood P.M., Nathans J.;
RT "Identification and characterization of a conserved family of protein serine/threonine phosphatases homologous to Drosophila retinal degeneration C.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:11639-11644(1997).

RN [2]
RP SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
RC Tissue-Fetal brain;
RX MEDLINE=97358589; PubMed=9215685;
RA Montini E., Ruggeri E.I., van de Vosse E., Andolfi G., Mariani M., Puccia A.A., Consales G.G., den Dunnen J.T., Ballabio A., Franco B.;
RT "A novel human serine-threonine phosphatase related to the Drosophila retinal degeneration C (rdgc) gene is selectively expressed in sensory neurons of neural crest origin.";
RL Hum. Mol. Genet. 6:1137-1145(1997).

RN [3]
RP SEQUENCE FROM N.A.
RC Tissue-Retina;
RX MEDLINE=98104127; PubMed=9430683;
RA Huang X., Honkanen R.E.;
RT "Molecular cloning, expression, and characterization of a novel human serine/threonine protein phosphatase, PP7, that is homologous to Drosophila retinal degeneration C gene product (rdgc).";
RL J. Biol. Chem. 273:1462-1468(1998).

RN [4]
RP SEQUENCE OF 1-355 FROM N.A.
RA Graham D.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.

RN [5]
RP Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RL Wray P.;
CC -1- FUNCTION: MAY HAVE A ROLE IN THE RECOVERY OR ADAPTATION RESPONSE OF PHOTORECEPTORS. MAY HAVE A ROLE IN DEVELOPMENT. MAXIMAL ACTIVITY IS OBSERVED AT PH 8.0.

CC -1- CATALYTIC ACTIVITY: A phosphoprotein + H(2)O = a protein + phosphate.

CC -1- COFACTOR: MAGNESIUM.

CC -1- ENZYME REGULATION: ACTIVATED BY CALCIUM.

CC -1- ALTERNATIVE PRODUCTS: 5 ISOFORMS: 1 (SHOWN HERE), 1A, 1B, 2 AND 3; ARE PRODUCED BY ALTERNATIVE SPLICING. ISOFORM 3 MAY HAVE NO FUNCTIONAL SIGNIFICANCE.

CC -1- TISSUE SPECIFICITY: DETECTED IN RETINA AND RETINAL DERIVED Y-79 RETINOBLASTOMA CELLS. ALSO FOUND IN FETAL BRAIN.

CC -1- SIMILARITY: BELONGS TO THE PPP FAMILY OF PHOSPHATASES.

CC -1- SIMILARITY: CONTAINS 1 IO DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: AF023455; AAB82795.1; -
DR EMBL: X97867; CA66461.1; -
DR EMBL: AF027977; AAC05825.1; -
DR EMBL: Z94056; CAB40074.1; -
DR EMBL: AL096700; CAB86407.1; -
DR HSP: P08129; 1PTM.
DR Genew: HSN29243; PPEF1.
DR MW: 300109; -
DR InterPro: IPR002048; EF-hand.
DR InterPro: IPR000048; I0-region.
DR InterPro: IPR004843; W-peptidase.

```

DR InterPro: IPR004844; S/T_phosphatase.
DR Pfam: PF00036; ehand; 3.
DR Pfam: PF00149; Metallophos; 1.
DR Pfam: PF00612; IQ; 1.
DR PRINTS: PR00114; STPHPTASE.
DR Prodom: PD000012; EF-hand; 1.
DR Prodom: PD000252; S/T_phosphatase; 2.
DR SMART: SM00054; Eph; 2.
DR SMART: SM00015; IQ; 1.
DR SMART: SM00156; PRAC; 1.
DR PROSITE: PS00018; EF_HAND; 2.
DR PROSITE: PS00125; SER_THR_PHOSPHATASE; 1.
DR PROSITE: PS50096; IQ; 1.
DR HydroLase: Calcium-binding; Magnesium; Iron; Manganese; Repeat;
KM Alternative splicing.
FT DOMAIN 18 43
FT DOMAIN 121 455
FT DOMAIN 496 507
FT CA_BIND 579 590
FT CA_BIND 619 630
FT METAL 172 172
FT METAL 174 174
FT METAL 201 201
FT METAL 233 233
FT ACT_SITE 234 234
FT METAL 285 285
FT METAL 403 403
FT VASPLIC 79 132
FT VASPLIC 328 355
FT VASPLIC 356 376
FT VASPLIC 377 653
FT VASPLIC 356 417
FT CONFLICT 367 367
SQ SEQUENCE 653 AA; 75792 MW; DFB78C444EE6484 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 2e+02; Length 653;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 640 RYEDL 644

RESULT 37
DNLJ_BACST STANDARD; PRT; 670 AA.
ID DNLJ_BACST
AC 087703;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA ligase (EC 6.5.1.2) [Polydeoxyribonucleotide synthase (NAD+)].
GN LIGA OR LIG.
OS Bacillus stearothermophilus.
OC Bacteria; Firmicutes; Bacillales; Geobacillus.
OX NCBI_Taxid=1422;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCB 1503;
RX MEDLINE=99337489; PubMed=10407164;
RA Brannigan J.A., Ashford S.R., Doherty A.J., Timson D.J., Wigley D.B.;
RT "Nucleotide sequence, heterologous expression and novel purification
of DNA ligase from Bacillus stearothermophilus.";
RL Biochim. Biophys. Acta 1432:413-418(1999).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS) OF 1-318.
RX MEDLINE=99148111; PubMed=10368271;
RA Singleton M.R., Hakansson K., Timson D.J., Wigley D.B.;
RT "Structure of the adenylation domain of an NAD+-dependent DNA
ligase.";
RL Structure 7:35-42(1999).
CC -!- FUNCTION: THIS PROTEIN CATALYZES THE FORMATION OF PHOSPHODIESTER

```

```

CC LINKAGES BETWEEN 5'-PHOSPHORYL AND 3'-HYDROXYL GROUPS IN DOUBLE-
CC STRANDED DNA USING NAD AS A COENZYME AND AS THE ENERGY SOURCE FOR
CC THE REACTION. IT IS ESSENTIAL FOR DNA REPLICATION AND REPAIR OF
CC DAMAGED DNA (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: NAD(+) + {deoxyribonucleotide}(N) +
CC {deoxyribonucleotide}(N) = AMP + nicotinamide nucleotide +
CC -!- SIMILARITY: BELONGS TO THE NAD-DEPENDENT DNA LIGASE FAMILY.
CC -!- SIMILARITY: CONTAINS 1 BCT DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ011676; CA09732.1; -.
DR PDB: 1B04; 22-NOV-99.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR004150; DNA_ligase_OR.
DR InterPro: IPR001679; DNA_ligase.
DR InterPro: IPR003583; HhH.1.
DR InterPro: IPR004445; HhH.
DR InterPro: IPR004149; Znf_DNA_ligase_C4.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00633; HhH; 1.
DR Pfam: PF01653; DNA_ligase_N; 1.
DR Pfam: PF03119; DNA_ligase_ZBD; 1.
DR Pfam: PF03120; DNA_ligase_OR; 1.
DR Prodom: PD003944; DNA_ligase; 1.
DR SMART: SM00292; BRCT; 1.
DR SMART: SM00276; HhH; 3.
DR SMART: SM00532; LIGANC; 1.
DR TIGRfam: TIGR00575; dnlj; 1.
DR PROSITE: PS0172; BRCT; 1.
DR PROSITE: PS01055; DNA_LIGASE_N1; 1.
DR PROSITE: PS01056; DNA_LIGASE_N2; 1.
DR Ligase; DNA repair; DNA replication; NAD; 3D-structure.
KM LIGA; 589 670
FT DOMAIN 114 114
FT BINDING 114 114
SQ SEQUENCE 670 AA; 74230 MW; B52462314CF9ACF5 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 2e+02; Length 670;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 5
DB 122 VYED 126

RESULT 38
YDOA_SCHPO STANDARD; PRT; 716 AA.
ID YDOA_SCHPO
AC 013730;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein C15A10.10 in chromosome I.
GN SPAC15A10.10.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_Taxid=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21648401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,

```

RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holtroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Moorey P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinovitch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Wolke G., Aert R., Robben J., Grympey B.,
RA Wellens J., Vansteels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehnach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mambrot R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Kochet M., Galliard C., Tallada V.A., Garzon A., Rhode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.,
RA "The genome sequence of Schizosaccharomyces pombe".
RL Nature 415:871-880(2002).
CC -1- SIMILARITY: CONTAINS 6 KELCH REPEATS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, 297208: CAB10107.1; -
DR InterPro: IPR001798; Kelch.
DR Pfam: PF01344; Kelch; 2.
KM Hypothetical protein, Repeat.
FT REPEAT 276 322
FT REPEAT 327 381
FT REPEAT 390 439
FT REPEAT 452 499
FT REPEAT 508 558
FT REPEAT 585 635
SQ SEQUENCE 716 AA: 83605 MW: BD2F50D06F5640D CRC64;
Query Match 87.1%; Score 27; DB 1; Length 716;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 VYED 5
DB 424 VYED 428

RL Virology 189:304-316(1992).
CC -1- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING.
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC HSV-1 56, EBV BRFL1, HCMV UL104, AND VAV 54.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, M86664: M802491.1; -
DR PIR: B36801; WZBE8.
DR InterPro: IPR002660; Herpes_UL6.
DR Pfam: PF01763; Herpes_UL6; 1.
DR Prodom: PD003210; Herpes_UL6; 1.
SQ SEQUENCE 753 AA: 83992 MW: C5E118F78BBD203 CRC64;
Query Match 87.1%; Score 27; DB 1; Length 753;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2 RYEDL 6
DB 108 RYEDL 112

RESULT 40
PARC_CAUCR STANDARD; PRT; 759 AA.
AC 054478:
DT 15-DEC-1998 (Rel. 37, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Topoisomerase IV subunit A (EC 5.99.1.1).
GN PARC OR C01566.
OS Caulobacter crescentus.
CC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter
OX NCB1_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=98086097; PubMed=9426128;
RA Ward D.V., Newton A.;
RT "Requirement of topoisomerase IV parC and parE genes for cell cycle
RT progression and developmental regulation in Caulobacter crescentus";
RL Mol. Microbiol. 26:897-910(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RA Ward D.V., Newton A.;
RL Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nieman W.C., Feldblyum T.V., Iwab M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Pirocka I., Nelson W.C., Newton A.S., Swinn M.L., Hatt D.E.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Khouri H., Shetty J., Berry K.,
RA Kolonay J.F., Salt J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utechtack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus".
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: TOPOISOMERASE IV IS ESSENTIAL FOR CHROMOSOME
CC SEGREGATION. IT HAS RELAXATION OF SUPERCOILED DNA ACTIVITY.
CC PERFORMS THE DECATENATION EVENTS REQUIRED DURING THE REPLICATION
CC OF A CIRCULAR DNA MOLECULE (BY SIMILARITY).


```

CC -1- SUBUNIT: COMPOSED OF TWO SUBUNITS: PARC AND PARF.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- SIMILARITY: STRONG, WITH THE A SUBUNIT OF GINASE.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U94696; AAC38043.1; -
DR EMBL: U86302; AAF14339.1; -
DR EMBL: AF005831; AAK23545.1; -
DR HSP: P09097; IAB4.
DR TIGR: CC1566; -
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV.1.
DR ProDom: PD000742; DNA_topoisomIV.1.
DR SMART: SM00434; TOP4C; 1.
DR TIGRfams: TIGR01062; parC_Gneg; 1.
DR TopoIsomerase: Isomerase; DNA-binding; Complete proteome.
FT ACT_SITE 132 132 DNA_CLEAVAGE (BY SIMILARITY).
FT CONFLICT 498 498 R -> A (IN REF. 1 AND 2).
SQ SEQUENCE 759 AA; 83521 MW; 65DDA34154A347D4 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 759;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
Db 408 VRYED 412

RESULT 41
DIVL CAUCR STANDARD; PRT; 769 AA.
ID DIVL CAUCR 09RQ09: Q9A2S2;
AC 09RQ09: Q9A2S2;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Sensor protein divL (EC 2.7.3.-).
GN DIVL OR CC3484.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
CC Caulobacter.
CC NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A. AND AUTOPHOSPHORYLATION SITE.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=10557274;
RA Wu J., Ohta N., Zhao J.L., Newton A.;
RA "A novel bacterial tyrosine kinase essential for cell division and
RA differentiation.";
RL Proc. Natl. Acad. Sci. U.S.A. 96:13068-13073(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeMay R.T., Dodson R.D., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utechtack T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RA "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
CC -1- FUNCTION: Required for cell division and growth. It catalyzes the
CC phosphorylation of CtrA and activates transcription in vitro of
CC the cell cycle-regulated flhF promoter.

```

```

CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- PTM: Autophosphorylated.
CC -1- SIMILARITY: CONTAINS 1 HISTIDINE KINASE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF083422; AAF08344.2; -
DR EMBL: AE006007; AAK25446.1; -
DR TIGR: CC3484; -
DR PhosSite: Q9RQ09; -
DR InterPro: IPR004358; Bact_sens-Pr_C.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR004359; His_Kin_sig.
DR InterPro: IPR000014; PAS_domain.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; signal; 1.
DR PRINTS: PR00344; BCTRLSENSOR.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00388; HisKA; 1.
DR TIGRfams: TIGR00229; sensory_box; 2.
DR PROSITE: PS50109; His_Kin; 1.
KW Sensory transduction; Transferase; Kinase; Transmembrane;
KW Phosphorylation; Complete proteome.
FT TRANSMEM 5 26 POTENTIAL.
FT DOMAIN 547 758 HISTIDINE KINASE.
FT DOMAIN 9 221 ALA-RICH.
FT MOD_RES 550 550 PHOSPHORYLATION (AUTO-).
FT CONFLICT 200 200 Q -> H (IN REF. 1).
FT CONFLICT 216 216 V -> E (IN REF. 1).
SQ SEQUENCE 769 AA; 82796 MW; 002B2428F18A57EF CRC64;

Query Match 87.1%; Score 27; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 351 RYEDL 355

RESULT 42
UL06 VZVD STANDARD; PRT; 769 AA.
ID UL06 VZVD AC P09302;
AC P09302;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-DEC-1992 (Rel. 24, Last annotation update)
DE Varion gene 54 protein.
GN 54.
OS Varicella-zoster virus (strain Dumas) (VZV).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicelloviruses.
CC NCBI_TaxID=10358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8630657; PubMed=3018124;
RA Davison A.J., Scott J.E.;
RA "The complete DNA sequence of varicella-zoster virus.";
RL J. Gen. Virol. 67:1759-1816(1986)
CC -1- FUNCTION: PRESUMED VIRION PROTEIN; POSSIBLE ROLE IN DNA
CC PACKAGING (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO FAMILY THAT GROUPS TOGETHER HSV-1 UL6,
CC EBV-1 56, EBV BBRF1, HCMV UL104, AND VZV 54.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```


CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X04370; CAA27937.1; -
DR PIR; B27215; W2BE54.
DR InterPro; IPR002660; Herpes_UL6.
DR Pfam; PF01763; Herpes_UL6; 1.
DR Prodom; PD003210; Herpes_UL6; 1.
SQ SEQUENCE 769 AA; 86780 MM; 5ABD7EDA6D783ECF CRC64;

Query Match 87.1%; Score 27; DB 1; Length 769;
Best Local Similarity 100.0%; Pred. No. 2,4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 119 RYEDL 123

RESULT 43
ORP5_MOUSE STANDARD; PRT; 874 AA.
AC Q9ER64; Q99NE5; 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Oxytelin binding protein-related protein 5 (OSBP-related protein 5)
DE (Orp-5) (Oxytelin-binding protein homologue 1).
GN OSBP5 OR OSBP2 OR OSBP1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=129/SV.
RA MEDLINE=20519229; Pubmed=11063728;
RA Engeman S., Stredlecke M., Paulsen M., Franck O., Reinhardt R.,
RA Lane N., Reik W., Walter J.;
RT "Sequence and functional comparison in the Beckwith-Wiedemann region;
RT implications for a novel imprinting centre and extended imprinting.";
RL Hum. Mol. Genet. 9:2691-2706(2000)
CC -1- SIMILARITY: BELONGS TO THE OSBP FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ278263; CAC16404.2; ALT_INIT.
DR EMBL; AJ276505; CAC27351.1; ALT_INIT.
DR MGD; MGI:1930265; Osbp2.
DR InterPro; IPR000648; Oxytelin_BP.
DR InterPro; IPR001849; PH.
DR Pfam; PF00169; PH; 1.
DR Pfam; PF01237; Oxytelin_BP; 1.
DR SMART; SM00233; PH; 1.
DR PROSITE; PS01013; OSBP; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Lipid transport; Transport.
FT DOMAIN 126 243 PH.
FT CONFLICT 37 44
SQ SEQUENCE 874 AA; 98921 MW; FBC41FA8E219F5E3 CRC64;
CAC27351).
Query Match 87.1%; Score 27; DB 1; Length 874;
Best Local Similarity 100.0%; Pred. No. 2,7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 707 RYEDL 711

RESULT 44
PKC1_ASPNG STANDARD; PRT; 1096 AA.
AC Q00078;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein kinase C-like (EC 2.7.1.-).
GN PKCA.
OS Aspergillus niger.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
OC Eutiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.
OX NCBI_TaxID=5061;
[1]
RN RP SEQUENCE FROM N.A.
RA MEDLINE=96158841; Pubmed=8569684;
RA Morawetz R., Lendenfeld T., Mischak H., Muehlbauer M., Gruber F.,
RA Goodnight J., de Graaf L.H., Visser J., Mushinski J.F.,
RA Kubicek C.P.;
RT "Cloning and characterisation of genes (pkc1 and pkca) encoding
RT protein kinase C homologues from Trichoderma reesei and Aspergillus
RT niger.";
RL Mol. Genet. 250:17-28(1996).
CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHOSPHO-ESTER AND DAG
CC BINDING DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
CC PKC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U10549; AAG97433.1; -
DR HSSP; Q63450; 1A06.
DR InterPro; IPR000008; C2.
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000719; Euk_pkinase.
DR InterPro; IPR000961; Pkinase_C.
DR InterPro; IPR000861; REM_repeat.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR Pfam; PF00130; DAG_PE-bind; 2.
DR Pfam; PF00433; Pkinase_C; 1.
DR Pfam; PF02185; HRI; 2.
DR Prodom; PD000001; Euk_pkinase; 1.
DR SMART; SM00109; C1; 2.
DR SMART; SM00239; C2; 1.
DR SMART; SM00074; HRI; 2.
DR SMART; SM00133; S_TR_X; 1.
DR SMART; SM00220; S_TR_X; 1.
DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
DR PROSITE; PS50061; DAG_PE_BIND_DOM_2; 2.
DR PROSITE; PS01017; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS01018; PROTEIN_KINASE_ST; 1.
KW transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
KW Phospho-ester binding; Repeat.
FT DOMAIN 460 507 PHOSPHO-ESTER AND DAG BINDING 1;
FT DOMAIN 528 577 PHOSPHO-ESTER AND DAG BINDING 2;
FT DOMAIN 771 1030 PROTEIN KINASE.
FT NP_BIND 777 785 ATP (BY SIMILARITY).
FT BINDING 800 800 ATP (BY SIMILARITY).

FT ACU_SITE 896 896 BY SIMILARITY
 SO SEQUENCE 1096 AA; 122234 MW; 859B2B35305B08 CRC64;
 Query Match 87.1%; Score 27; DB 1; Length 1096;
 Best Local Similarity 100.0%; Pred. No. 3.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 218 RYEDL 222

RESULT 45
 RPOB_THEAO STANDARD; PRT; 1119 AA.

AC 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE DNA-directed RNA polymerase beta chain (EC 2.7.7.6) (Transcriptase
 beta chain) (RNA polymerase beta subunit).
 GN RPOB.
 OS Thermus aquaticus.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Thermales;
 OC Thermaceae; Thermus.
 OX NCBI_TaxID=271;
 RN [1]
 RP SEQUENCE FROM N.A. AND X-RAY CRYSTALLOGRAPHY (3.3 ANGSTROMS).
 RX MEDLINE=99428144; PubMed=10499798;
 RA Zhang G., Campbell E.A., Minkhin L., Richter C., Severinov K.,
 RA Darts S.A.;
 RT "Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 A
 resolution.";
 RL Cell 98:811-824(1999).
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
 OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
 SUBSTRATES.
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
 (RNA)(N).
 CC -1- SUBUNIT: THE ENZYME CONSISTS OF THE SIGMA CHAIN AND THE CORE
 ENZYME WHICH IS COMPOSED OF 2 ALPHA CHAINS, 1 BETA CHAIN, AND 1
 BETA' CHAIN.
 CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA CHAIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL, Y19223; CAB65465.2; -
 DR PDB; 1HQW; 07-FEB-01.
 DR InterPro; IPR001572; RNA_POL_B.
 DR Pfam; PF00562; RNA_POL_B; 2.
 DR PROSITE; PS01166; RNA_POL_BETA; 1.
 KW Transferase; Transcription; DNA-directed RNA polymerase;
 KW 3D-structure; 1119 AA; 124757 MW; 2CF66FA9C77F33 CRC64;
 SQ SEQUENCE

Query Match 87.1%; Score 27; DB 1; Length 1119;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYEDL 5
 DB 613 VYEDL 617

RESULT 46
 KPCL_COCHHE STANDARD; PRT; 1174 AA.

AC 042632;
 DT 15-JUL-1999 (Rel. 38, Created)
 DT 15-JUL-1999 (Rel. 38, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Protein kinase C-like (EC 2.7.1.1).
 GN PKC1.
 OS Cochliobolus heterostrophus (Drechslera maydis).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;
 OC Pleosporales; Pleosporaceae; Cochliobolus.
 OX NCBI_TaxID=5016;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 48329 / C2;
 RA Oeser B.M., Yoder O.C.;
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 ZINC-DEPENDENT PHORBOL-ESTER AND DAG
 BINDING DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 CC PKC SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; Y15839; CAA75801.1; -
 DR HSSP; 063450; 1A06.
 DR InterPro; IPR000008; C2.
 DR InterPro; IPR002219; DAG_PE_BIND.
 DR InterPro; IPR000719; Euk_Pkinase.
 DR InterPro; IPR000961; Pkinase_C.
 DR InterPro; IPR000861; REM_repeat.
 DR InterPro; IPR002290; Ser_thr_Pkinase.
 DR Pfam; PF00069; Pkinase; 1.
 DR Pfam; PF00130; DAG-pe-bind; 2.
 DR Pfam; PF00433; Pkinase_C; 1.
 DR Pfam; PF02185; HRI; 2.
 DR ProDom; PD000001; Euk_Pkinase; 1.
 DR SMART; SM00109; C1; 2.
 DR SMART; SM00239; C2; 1.
 DR SMART; SM00074; HRI; 2.
 DR SMART; SM00133; S_TK_X; 1.
 DR SMART; SM00220; S_TK_X; 1.
 DR PROSITE; PS00479; DAG_PE_BIND_DOM_1; 2.
 DR PROSITE; PS00081; DAG_PE_BIND_DOM_2; 2.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Zinc;
 KW Phorbol-ester binding; Repeat.

KM PHORBOL-ESTER AND DAG BINDING 1.
 KM PHORBOL-ESTER AND DAG BINDING 2.
 FT DOMAIN 527 576
 FT DOMAIN 704 709
 FT DOMAIN 785 792
 FT DOMAIN 849 1108
 FT NP_BIND 855 863
 FT BINDING 878 878
 FT ACT_SITE 974 974
 SQ SEQUENCE 1174 AA; 130506 MW; 26A4ADDA2849F37C CRC64;

Query Match 87.1%; Score 27; DB 1; Length 1174;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 219 RYEDL 223

RESULT 47

ID	PROT	SEQUENCE	STANDARD	CHAIN	PROT	SEQUENCE	STANDARD	CHAIN	ID
1	1393	AA.			1	1393	AA.		1
2	1393	AA.			2	1393	AA.		2
3	1393	AA.			3	1393	AA.		3
4	1393	AA.			4	1393	AA.		4
5	1393	AA.			5	1393	AA.		5
6	1393	AA.			6	1393	AA.		6
7	1393	AA.			7	1393	AA.		7
8	1393	AA.			8	1393	AA.		8
9	1393	AA.			9	1393	AA.		9
10	1393	AA.			10	1393	AA.		10
11	1393	AA.			11	1393	AA.		11
12	1393	AA.			12	1393	AA.		12
13	1393	AA.			13	1393	AA.		13
14	1393	AA.			14	1393	AA.		14
15	1393	AA.			15	1393	AA.		15
16	1393	AA.			16	1393	AA.		16
17	1393	AA.			17	1393	AA.		17
18	1393	AA.			18	1393	AA.		18
19	1393	AA.			19	1393	AA.		19
20	1393	AA.			20	1393	AA.		20
21	1393	AA.			21	1393	AA.		21
22	1393	AA.			22	1393	AA.		22
23	1393	AA.			23	1393	AA.		23
24	1393	AA.			24	1393	AA.		24
25	1393	AA.			25	1393	AA.		25
26	1393	AA.			26	1393	AA.		26
27	1393	AA.			27	1393	AA.		27
28	1393	AA.			28	1393	AA.		28
29	1393	AA.			29	1393	AA.		29
30	1393	AA.			30	1393	AA.		30
31	1393	AA.			31	1393	AA.		31
32	1393	AA.			32	1393	AA.		32
33	1393	AA.			33	1393	AA.		33
34	1393	AA.			34	1393	AA.		34
35	1393	AA.			35	1393	AA.		35
36	1393	AA.			36	1393	AA.		36
37	1393	AA.			37	1393	AA.		37
38	1393	AA.			38	1393	AA.		38
39	1393	AA.			39	1393	AA.		39
40	1393	AA.			40	1393	AA.		40
41	1393	AA.			41	1393	AA.		41
42	1393	AA.			42	1393	AA.		42
43	1393	AA.			43	1393	AA.		43
44	1393	AA.			44	1393	AA.		44
45	1393	AA.			45	1393	AA.		45
46	1393	AA.			46	1393	AA.		46
47	1393	AA.			47	1393	AA.		47
48	1393	AA.			48	1393	AA.		48
49	1393	AA.			49	1393	AA.		49
50	1393	AA.			50	1393	AA.		50
51	1393	AA.			51	1393	AA.		51
52	1393	AA.			52	1393	AA.		52
53	1393	AA.			53	1393	AA.		53
54	1393	AA.			54	1393	AA.		54
55	1393	AA.			55	1393	AA.		55
56	1393	AA.			56	1393	AA.		56
57	1393	AA.			57	1393	AA.		

	Query Match	Score 27; DB 1; Length 1393;
	Best Local Similarity	66.7%; Pred. No. 4.5e+02;
Matches	4; Conservative	2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
::|||
Db 1048 IXYEDL 1053

RESULT 48
VTFL_FUNHE STANDARD; PRT, 1704 AA.

ID VTFL_FUNHE AC Q90508;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vitellogenin I precursor (VTE I) [Contains: Lipovitellin 1 (LV1); Phosvitin (PV); Lipovitellin 2 (LV2)].
GN VTGI.
OS Fundulus heteroclitus (Killifish) (Mummichog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorphia; Atherinomorpha; Cyprinodontiformes; Fundulidae; Fundulus.
NCBI_TaxID=8078;
[1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1220-1225.
RC TISSUE=Liver;
RX MEDLINE=96029150; PubMed=7563139;
RA Latteur G.J. Jr., Byrne B.M., Kaugungo J., Nelson L.D., Greenberg R.M., Wallace R.A.;
RT "Fundulus heteroclitus vitelligenin: the deduced primary structure of a piscine precursor to noncrystalline, liquid-phase yolk protein.";
RL J. Mol. Evol. 41:505-521(1995).
CC -1- FUNCTION: PRECURSOR OF THE MAJOR EGG-YOLK PROTEINS THAT ARE SOURCES OF NUTRIENTS DURING EARLY DEVELOPMENT OF OVIPAROUS ORGANISMS.
CC -1- TISSUE SPECIFICITY: PRODUCED BY THE LIVER, SECRETED INTO THE BLOOD AND THEN SEQUESTERED BY RECEPTOR MEDIATED ENDOCYTOSIS INTO GROWING OOCYTES, WHERE IT IS GENERALLY CLEAVED, GIVING RISE TO THE RESPECTIVE YOLK COMPONENTS COMPOSED OF COMPLEX SITE OF SMALL CLEAVAGE PRODUCTS.
CC -1- PTM: PHOSVITIN, AN EGG YOLK STORAGE PROTEIN, IS ONE OF THE MOST HIGHLY PHOSPHORYLATED (10x) PROTEINS IN NATURE (BY SIMILARITY).
CC -1- PTM: THE N-TERMINAL OF THE BLOOD VITELLOGENIN IS BLOCKED.
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announcement/or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U07055; AAA93123.1; -;
DR InterPro: IPR001747; Lipid.transprt_N.
DR InterPro: IPR001846; VMF_D.
DR Pfam: PF00094; vwd; 1.
DR Pfam: PF01347; Vitellogenin_N; 1.
KW SMART: SM00216; WMD; 1.
RW Glycoprotein; Phosphorylation; Storage protein; Signal.
FT SIGNAL 1 14 POTENTIAL.
FT CHAIN 15 1704 VITELLOGENIN I.
FT CHAIN 15 ? LIPOVITELLIN 1.
FT CHAIN ? ? PHOSVITIN.
FT CHAIN ? 1704 LIPOVITELLIN 2.
FT DOMAIN 1081 1205 SER-RICH.
FT POLY-AA.
FT CARBOHYD 446 446 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 635 635 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 903 903 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 908 908 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1019 1019 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1080 1080 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1121 1121 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1174 1174 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1285 1285 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1322 1322 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1375 1375 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1379 1379 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1405 1405 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1456 1456 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 1512 1512 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SO SEQUENCE 1704 AA; 167954 MW; 557A82804611221 CRC64;

Query Match 87.1%; Score 27; DB 1; Length 1704;
 Best Local Similarity 100.0%; Pred. No. 5.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 342 RYEDL 346

RESULT 49
 ID UTRO_HUMAN STANDARD; PRT; 3433 AA.
 AC P46939;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE utrophin (Dystrophin-related protein 1) (DRP1) (DRP).
 GN UTRO OR DMDL.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Tinsley J.M., Blake D.J., Roche A., Fairbrother U., Riss J.,
 RA Byth B.C., Knight A.E., Kendrick-Jones J., Suthers G.K., Love D.R.,
 RA Edwards Y.H., Davies K.E.;
 RT "Primary structure of dystrophin-related protein.";
 RL Nature 360:591-593(1992).
 RN [2]
 RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 144-261.
 RX MEDLINE=99141377; PubMed=8807274;
 RA Keep N.H., Norwood F.L.M., Moores C.A., Winder S.J.,
 RA Kendrick-Jones J.;
 RT "The 2.0-A structure of the second calponin homology domain from the
 RT actin-binding region of the dystrophin homologue utrophin.";
 RL J. Mol. Biol. 285:1257-1264(1999).
 RN [3]
 RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 31-256.
 RX MEDLINE=20113481; PubMed=10647184;
 RA Keep N.H., Winder S.J., Moores C.A., Walke S., Norwood F.L.M.,
 RA Kendrick-Jones J.;
 RT "Crystal structure of the actin-binding region of utrophin reveals a
 RT head-to-tail dimer.";
 RL Structure 7:1539-1546(1999).
 CC -1- FUNCTION: MAY PLAY A ROLE IN ANCHORING THE CYTOSKELETON TO THE
 CC PLASMA MEMBRANE (BY SIMILARITY TO DYSTROPHIN).
 CC -1- SUBCELLULAR LOCATION: NEUROMUSCULAR JUNCTION.
 CC -1- TISSUE SPECIFICITY: MUSCLE.
 CC -1- SIMILARITY: STRONG, TO DYSTROPHIN.
 CC -1- SIMILARITY: THE ACTIN-BINDING DOMAIN IS OF A TYPE FOUND IN MANY
 CC ACTIN-BINDING PROTEINS (SUCH AS ACTININ, DYSTROPHIN, FILAMIN,
 CC ABP-120, ABP-180, OR BETA-FODRIN).
 CC -1- SIMILARITY: CONTAINS 2 CALPONIN-HOMOLOGY (CH) DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 WW DOMAIN.
 CC -1- SIMILARITY: CONTAINS 20 SPECTRIN REPEATS.
 CC -1- SIMILARITY: CONTAINS 1 Z2-TYPE ZINC FINGER.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL: X69086; CAA48829.1; -
 CC FIR: S28381; S28381.
 CC PDB: 1BHD; 16-FEB-99.
 CC PDB: 1OAG; 01-JAN-00.
 CC Genew; HGNC:12635; UTRN.
 DR MIM: 128240; -
 DR InterPro: IPR001589; Actbind_actnin.
 DR InterPro: IPR001715; Calponin-like.
 DR InterPro: IPR002017; Spectrin.
 DR InterPro: IPR003132; TatrH.
 DR InterPro: IPR002349; WW.
 DR InterPro: IPR001202; WW_Rsp5_WMP.
 DR InterPro: IPR000453; Znf_ZZ.
 DR Pfam: PF00307; CH; 2.
 DR Pfam: PF00435; Spectrin; 19.
 DR Pfam: PF00569; Z2; 1.
 DR PRINTS: PRO0403; WWDOMAIN.
 DR SMART: SM00033; CH; 2.
 DR SMART: SM00150; SPEC; 18.
 DR SMART: SM00319; TatrH; 1.
 DR SMART: SM00456; WW; 1.
 DR SMART: SM00291; ZNF_ZZ; 1.
 DR PROSITE: PS00019; ACTININ_1; 1.
 DR PROSITE: PS00020; ACTININ_2; 1.
 DR PROSITE: PS00159; WW_DOMAIN_1; 1.
 DR PROSITE: PS00020; WW_DOMAIN_2; 1.
 DR PROSITE: PS01357; ZF_ZZ_1; 1.
 DR PROSITE: PS01357; ZF_ZZ_2; 1.
 DR PROSITE: PS01357; ZF_ZZ_3; 1.
 KW Structural protein; Actin-binding; Calcium-binding; Cytoskeleton;
 KW Repeat; 3D-structure; Zinc-finger.
 FT DOMAIN 1 246
 FT 31 135
 FT 150 252
 FT 253 308
 FT 309 417
 FT 418 526
 FT 541 637
 FT 687 738
 FT 798 902
 FT 1016 1083
 FT 1125 1230
 FT 1248 1334
 FT 1432 1541
 FT 1544 1649
 FT 1652 1753
 FT 1910 1968
 FT 1976 2081
 FT 2258 2333
 FT 2399 2440
 FT 2443 2556
 FT 2559 2636
 FT 2658 2688
 FT 2691 2797
 FT 2812 2845
 FT 3064 3111
 FT ZN_FING 3064 3111
 SO SEQUENCE 3433 AA; 394488 MW; EAE9DBA09F858E5B CRC64;

Query Match 87.1%; Score 27; DB 1; Length 3433;
 Best Local Similarity 100.0%; Pred. No. 1.2e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
 DB 1317 RYEDL 1321

RESULT 50
DYH9_HUMAN STANDARD; PRT; 4486 AA.
ID Q9NTC9; OSNO28; 095494;
DI 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DI 16-OCT-2001 (Rel. 40; Last annotation update)
DE Ciliary dynein heavy chain (Axonemal dynein heavy chain) (Dynein heavy chain 9)
GN DNAB9 OR DNAB17L OR DNAB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Reed W., Moats-Staats B.M., Carson J.L., Leigh M.W., Collier A.M.;
RT "A ciliary dynein heavy chain whose expression is upregulated in differentiating airway epithelium.";
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Nasal epithelium;
RA Bartoloni L., Blouin J.L., Maiti A., Sainsbury A., Rossier C., Gehrig C., She J.X., Marron M.P., Lander E., Meeks M., Chung E., Jorissen M., Scott H.S., Delozier Blanchet C.D., Gardiner M., Antonarakis S.E.;
RT "Axonemal beta heavy chain dynein DNAB9: cDNA sequence, genomic structure and investigation of its role in primary ciliary dyskinesia.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 1874-1974 FROM N.A.
RC TISSUE=Nasal polyps;
RA Maiti A.X., Maitel M.G., Jorissen M., Volz A., Ziegler A., Bouvagnet P.;
RT "Chromosomal localization of human dynein heavy chain genes.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP CHARACTERIZATION.
RX MEDLINE=20558134; PubMed=1104725;
RA Reed W., Carson J.L., Moats-Staats B.M., Lucier T., Hu P.C., Brighton L., Gamblin T.M., Huang C.H., Leigh M.W., Collier A.M.;
RT "Characterization of an axonemal dynein heavy chain expressed early in airway epithelial ciliogenesis.";
RL Am. J. Respir. Cell Mol. Biol. 23:734-741(2000).
CC -1- FUNCTION: FORCE GENERATING PROTEIN OF RESPIRATORY CILIA. PRODUCES FORCE TOWARDS THE MINUS ENDS OF MICROTUBULES. DYNEIN HAS ATPASE ACTIVITY.
CC -1- SUBUNIT: CONSISTS OF AT LEAST TWO HEAVY CHAINS AND A NUMBER OF INTERMEDIATE AND LIGHT CHAINS.
CC -1- SIMILARITY: BELONGS TO THE DYNEIN HEAVY CHAIN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF257737; AAF69004.1; -;
DR EMBL: AJ404468; CAB94756.1; -;
DR EMBL: AJ132088; CA10561.1; -;
DR Gene: HGNC:2953; DNAB9.
DR MIM: 603330; -;
DR InterPro: IPR004273; Dynein_heavy.
DR Pfam: PF03028; Dynein_heavy; 1.
KW Motor protein; Microtubules; Dynein; ATP-binding; Coiled coil.
FT DOMAIN 381 410 COILED COIL (POTENTIAL).
FT DOMAIN 504 529 COILED COIL (POTENTIAL).

FT	DOMAIN	639	662	COILED COIL (POTENTIAL).
FT	DOMAIN	752	823	COILED COIL (POTENTIAL).
FT	DOMAIN	1326	1355	COILED COIL (POTENTIAL).
FT	DOMAIN	3051	3154	COILED COIL (POTENTIAL).
FT	DOMAIN	3285	3341	COILED COIL (POTENTIAL).
FT	DOMAIN	3640	3675	COILED COIL (POTENTIAL).
FT	NP_BIND	1870	1877	ATP (POTENTIAL).
FT	NP_BIND	2151	2158	ATP (POTENTIAL).
FT	NP_BIND	2478	2485	ATP (POTENTIAL).
FT	NP_BIND	2825	2832	ATP (POTENTIAL).
FT	CONFLICT	2505	2505	L -> V (IN REF. 2).
FT	CONFLICT	3678	3678	T -> A (IN REF. 2).
FT	CONFLICT	4374	4374	I -> M (IN REF. 2).
SQ	SEQUENCE	4486 AA;	511927 MW;	996DEFFDEB0B3EB1 CRC64;

Query Match
Best Local Similarity 87.1%; Score 27; DB 1; Length 4486;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 2 RYEDL 6
DB 258 RYEDL 262

Search completed: February 20, 2003, 13:32:38
Job time : 16.5714 secs

This Page Blank (uspto)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

CM protein - protein search, using sw model

Run on: February 20, 2003, 13:29:11 ; Search time 8.57143 Seconds

(without alignments)
144.233 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 1000 summaries

Database :

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mmc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	114	11 Q9D0K5	Q9D0K5 mus musculus
2	31	100.0	183	5 Q9VMC2	Q9VMC2 drosophila
3	31	100.0	281	17 Q980M6	Q980M6 sulfolobus
4	31	100.0	331	12 Q11316	Q11316 molluscum c
5	31	100.0	386	4 Q9Y5R3	Q9Y5R3 mus sapien
6	31	100.0	395	11 Q9Q0P4	Q9Q0P4 mus musculu
7	31	100.0	411	4 Q43916	Q43916 homo sapien
8	31	100.0	411	11 Q9E0C0	Q9E0C0 mus musculu
9	31	100.0	428	16 Q9RVQ3	Q9RVQ3 mus musculu
10	31	100.0	484	4 Q90ED5	Q90ED5 delinococcus
11	31	100.0	486	5 Q95TN9	Q95TN9 homo sapien
12	31	100.0	530	4 Q9Y6F2	Q9Y6F2 drosophila
13	31	100.0	530	11 Q88276	Q88276 mus musculu
14	31	100.0	531	4 Q9Y4C5	Q9Y4C5 homo sapien
15	31	100.0	597	5 Q9Y156	Q9Y156 drosophila
16	31	100.0	631	16 Q25212	Q25212 helicobacte

17	31	100.0	649	5 Q9V864	Q9V864 drosophila
18	31	100.0	748	12 Q98200	Q98200 molluscum c
19	31	100.0	973	5 Q45412	Q45412 caenorhabdi
20	31	100.0	6260	2 Q54299	Q54299 streptomyce
21	30	96.8	194	17 Q82TB3	Q82TB3 pyrobaculum
22	30	96.8	199	11 Q9D806	Q9D806 mus musculu
23	30	96.8	363	5 Q9VMC4	Q9VMC4 drosophila
24	30	96.8	540	17 Q8T0C3	Q8T0C3 methanosarc
25	30	96.8	579	16 Q9A7T6	Q9A7T6 caulobacter
26	30	96.8	595	5 Q8SRF6	Q8SRF6 encephalito
27	30	96.8	1322	10 Q23559	Q23559 arbidopsis
28	29	93.5	234	16 Q9A994	Q9A994 arbidopsis
29	29	93.5	247	12 Q9JFS0	Q9JFS0 caulobacter
30	29	93.5	303	16 Q93UE6	Q93UE6 ectromelia
31	29	93.5	355	16 Q926E4	Q926E4 rhizobium m
32	29	93.5	393	10 Q81244	Q81244 cucumis mel
33	29	93.5	405	10 Q22818	Q22818 arbidopsis
34	28	90.3	154	2 P96458	P96458 streptomyce
35	28	90.3	195	17 Q28741	Q28741 archaeoglob
36	28	90.3	228	17 Q976N0	Q976N0 sulfolobus
37	28	90.3	233	16 Q9CEP1	Q9CEP1 lactococcus
38	28	90.3	273	2 Q9RC18	Q9RC18 bacillus sp
39	28	90.3	273	12 Q9YRQ1	Q9YRQ1 atelina her
40	28	90.3	278	2 Q9ET84	Q9ET84 bacillus am
41	28	90.3	306	16 Q91732	Q91732 pseudomonas
42	28	90.3	333	10 Q82330	Q82330 arbidopsis
43	28	90.3	346	5 Q45982	Q45982 caenorhabdi
44	28	90.3	354	10 Q9FL45	Q9FL45 arbidopsis
45	28	90.3	359	10 Q940P4	Q940P4 arbidopsis
46	28	90.3	388	11 Q9R111	Q9R111 mus musculu
47	28	90.3	388	11 Q9WUE5	Q9WUE5 mus musculu
48	28	90.3	391	17 Q97B92	Q97B92 thermoplas
49	28	90.3	458	17 Q9Y129	Q9Y129 pyrococcus
50	28	90.3	472	2 Q93N59	Q93N59 coxiella bu
51	28	90.3	472	11 Q88199	Q88199 mus musculu
52	28	90.3	474	11 Q902L2	Q902L2 ratus norv
53	28	90.3	479	4 Q75099	Q75099 homo sapien
54	28	90.3	484	11 Q99NB0	Q99NB0 mus musculu
55	28	90.3	484	11 Q9EP78	Q9EP78 mus musculu
56	28	90.3	486	4 Q9NS84	Q9NS84 homo sapien
57	28	90.3	486	4 Q75667	Q75667 homo sapien
58	28	90.3	503	17 Q980D7	Q980D7 sulfolobus
59	28	90.3	507	16 Q99WJ7	Q99WJ7 staphylococ
60	28	90.3	541	16 Q98308	Q98308 rhizobium l
61	28	90.3	612	10 Q9SD68	Q9SD68 arbidopsis
62	28	90.3	622	17 Q93736	Q93736 pyrococcus
63	28	90.3	672	16 Q9P8S8	Q9P8S8 streptomyce
64	28	90.3	697	13 Q90YX2	Q90YX2 discoglossu
65	28	90.3	718	12 Q91BW8	Q91BW8 turkey herp
66	28	90.3	722	12 Q9E6R0	Q9E6R0 turkey herp
67	28	90.3	788	12 Q9C7D7	Q9C7D7 marek's dis
68	28	90.3	793	5 Q9VBX0	Q9VBX0 drosophila
69	28	90.3	828	5 Q8SX93	Q8SX93 drosophila
70	28	90.3	875	5 Q9N396	Q9N396 caenorhabdi
71	28	90.3	943	5 Q9VE63	Q9VE63 drosophila
72	28	90.3	1078	17 Q8TPH6	Q8TPH6 methanosarc
73	28	90.3	1107	16 Q8XDP6	Q8XDP6 escherichia
74	28	90.3	1658	10 Q8S622	Q8S622 cryza sativ
75	28	90.3	2258	12 Q89338	Q89338 human parat
76	27	87.1	52	2 Q9RGS3	Q9RGS3 staphylococ
77	27	87.1	73	16 Q9KA06	Q9KA06 bacillus ha
78	27	87.1	108	17 Q9YEP8	Q9YEP8 aeropyrum p
79	27	87.1	112	2 Q9L4E2	Q9L4E2 uncultured
80	27	87.1	112	2 Q9L4E1	Q9L4E1 uncultured
81	27	87.1	112	2 Q9L479	Q9L479 uncultured
82	27	87.1	127	3 Q9P895	Q9P895 emericella
83	27	87.1	128	5 Q9U683	Q9U683 culic pipie
84	27	87.1	133	17 Q28799	Q28799 archaeoglob
85	27	87.1	149	5 Q9U019	Q9U019 plasmodiob
86	27	87.1	157	2 Q9L419	Q9L419 corynebacte
87	27	87.1	163	2 Q9EUV9	Q9EUV9 klebsiella
88	27	87.1	163	2 Q9EUV8	Q9EUV8 klebsiella
89	27	87.1	163	2 Q9EXR7	Q9EXR7 klebsiella

90	27	87.1	163	2	Q9EXT6	Q9ext6 klebsiella	163	27	87.1	504	16	Q9ALD7	Q9alld7 rhizobium m
91	27	87.1	163	2	Q9EXT5	Q9ext5 klebsiella	164	27	87.1	516	2	Q9AN11	Q9an11 bradyrhizob
92	27	87.1	163	2	Q9EXT4	Q9ext4 klebsiella	165	27	87.1	516	2	Q9JHS0	Q9jhs0 bradyrhizob
93	27	87.1	163	2	Q9EXT3	Q9ext3 klebsiella	166	27	87.1	519	2	Q9JN44	Q9jn44 coxiella bu
94	27	87.1	181	17	Q8SZ0	Q8sz0 pyrobaculum	167	27	87.1	520	5	Q45411	Q45411 caenorhabdi
95	27	87.1	188	4	Q960B9	Q96qb9 homo sapien	168	27	87.1	522	2	Q9KH11	Q9kh11 streptomyces
96	27	87.1	189	17	Q96Y3	Q96y3 sulfobus	169	27	87.1	525	16	Q9HUA5	Q9hu5 pseudomonas
97	27	87.1	200	3	Q9HDG9	Q9hdg9 cryptococcu	170	27	87.1	527	17	Q29232	Q29232 archaeoglob
98	27	87.1	200	3	Q9HGG5	Q9hgg5 cryptococcu	171	27	87.1	543	2	Q9K2F3	Q9k2f3 uncultured
99	27	87.1	200	3	Q9HGG5	Q9hgg5 cryptococcu	172	27	87.1	550	3	Q9P8P1	Q9p8p1 cryptococcu
100	27	87.1	200	3	Q9HGG4	Q9hgg4 cryptococcu	173	27	87.1	550	4	Q9MTN2	Q9mtn2 homo sapien
101	27	87.1	200	16	Q922M8	Q922m8 rhizobium m	174	27	87.1	551	4	Q9S678	Q9s678 homo sapien
102	27	87.1	203	16	Q9ZHC2	Q9zhc2 streptococc	175	27	87.1	551	4	Q9NSA9	Q9nsa9 homo sapien
103	27	87.1	203	16	Q9ZHC2	Q9zhc2 streptococc	176	27	87.1	554	3	Q9P399	Q9p399 neurospora
104	27	87.1	225	16	Q66072	Q66072 canine herp	177	27	87.1	554	5	Q8SR05	Q8sr05 encephalito
105	27	87.1	225	16	Q66755	Q66755 aquifex neo	178	27	87.1	557	10	Q9XIC5	Q9xics arabidopsis
106	27	87.1	236	10	Q65802	Q65802 ceratopter	179	27	87.1	561	16	Q8U7S2	Q8u7s2 agrobacteri
107	27	87.1	239	10	Q8TQ65	Q8tg65 methanosarc	180	27	87.1	584	10	Q9AV27	Q9av27 cyzsa saliv
108	27	87.1	241	10	Q9FKD6	Q9fk65 arabidopsis	181	27	87.1	592	3	Q59900	Q59900 cryptococcu
109	27	87.1	245	2	Q51405	Q51405 pseudomonas	182	27	87.1	592	16	Q9NH21	Q9nq27 homo sapien
110	27	87.1	253	7	P79522	P79522 homo sapien	183	27	87.1	604	4	Q9NQ27	Q9nq27 rhizobium 1
111	27	87.1	255	16	Q99R22	Q99r22 staphylococ	184	27	87.1	605	4	Q8TB00	Q8tbu0 homo sapien
112	27	87.1	275	16	Q8ZMR1	Q8zmr1 salmonella	185	27	87.1	605	16	Q8XAK0	Q8xak0 escherichia
113	27	87.1	279	2	Q51406	Q51406 pseudomonas	186	27	87.1	606	11	Q9ETJ0	Q9etj0 mus musculu
114	27	87.1	283	16	Q9JUK4	Q9juk4 neisseria m	187	27	87.1	617	16	Q92206	Q92206 bacillus su
115	27	87.1	284	12	Q66105	Q66105 citrus leaf	188	27	87.1	624	17	Q8U410	Q8u410 pyrococcus
116	27	87.1	298	10	Q9LV42	Q9lv42 arabidopsis	189	27	87.1	630	16	Q8UER9	Q8uer9 agrobacteri
117	27	87.1	307	16	Q92VG4	Q92vg4 rhizobium m	190	27	87.1	631	16	Q9ZM07	Q9zm07 helicobacter
118	27	87.1	308	4	Q9BWA0	Q9bwa0 homo sapien	191	27	87.1	653	12	Q64906	Q64906 acetalaphne
119	27	87.1	312	17	Q9YD31	Q9y31 aeropyrum p	192	27	87.1	666	16	Q8Y828	Q8y828 listeria mo
120	27	87.1	315	5	Q9N3F7	Q9n3f7 caenorhabdi	193	27	87.1	685	16	Q92TJ0	Q92tj0 rhizobium m
121	27	87.1	324	16	Q9BLQ4	Q9blq4 rhizobium 1	194	27	87.1	693	2	Q9L8Z2	Q9l8z2 enterococcu
122	27	87.1	325	5	Q9VN98	Q9vn98 drosophila	195	27	87.1	693	4	Q9UPU1	Q9upu1 homo sapien
123	27	87.1	334	16	Q9K689	Q9k689 bacillus ha	196	27	87.1	694	16	Q926R0	Q926r0 chlamydia p
124	27	87.1	338	12	Q9MWZ1	Q9mwz1 dioscorea a	197	27	87.1	706	5	Q9Y417	Q9y417 drosophila
125	27	87.1	351	2	Q52572	Q52572 amycolatops	198	27	87.1	713	10	Q9FT72	Q9ft72 arabidopsis
126	27	87.1	352	16	Q9K9N7	Q9k9n7 bacillus ha	199	27	87.1	747	12	Q93296	Q93296 equine herp
127	27	87.1	355	2	Q51154	Q51154 neisseria m	200	27	87.1	769	2	Q9R009	Q9r009 caulobacter
128	27	87.1	359	2	Q9F016	Q9f016 xanthomonas	201	27	87.1	773	16	Q9A2S2	Q9a2s2 caulobacter
129	27	87.1	369	2	Q9L307	Q9l307 streptomyces	202	27	87.1	785	17	Q9YCC8	Q9ycc8 aeropyrum p
130	27	87.1	378	2	Q9R8V5	Q9r8v5 pseudomonas	203	27	87.1	809	16	Q9A704	Q9a704 caulobacter
131	27	87.1	378	12	Q9LNU9	Q9l9u9 cydia pomon	204	27	87.1	850	4	Q9ULF7	Q9ulf7 homo sapien
132	27	87.1	379	2	Q68541	Q68541 cellvibrio	205	27	87.1	874	11	Q8R510	Q8r510 mus musculu
133	27	87.1	380	13	Q90XHO	Q90xh0 brachydanio	206	27	87.1	882	16	Q92JW2	Q92jw2 rhizobium m
134	27	87.1	381	11	Q99MH7	Q99mh7 mus musculu	207	27	87.1	882	16	Q92JW2	Q92jw2 rhizobium m
135	27	87.1	389	16	Q9K464	Q9k464 streptomyces	208	27	87.1	891	3	Q9US77	Q9us77 schizosacch
136	27	87.1	393	13	Q9PTM7	Q9ptm7 brachydanio	209	27	87.1	926	10	Q81R05	Q81r05 arabidopsis
137	27	87.1	393	13	Q9DDE0	Q9dde0 brachydanio	210	27	87.1	976	16	Q9ANM4	Q9anm4 caulobacter
138	27	87.1	395	4	Q9GZX3	Q9gzx3 homo sapien	211	27	87.1	977	16	Q9BNV6	Q9bnv6 rhizobium 1
139	27	87.1	399	11	Q9DBO5	Q9db95 mus musculu	212	27	87.1	991	3	Q9E6V6	Q9e6v6 tuber magna
140	27	87.1	411	4	Q9GZS9	Q9gz95 homo sapien	213	27	87.1	1004	3	Q9HE23	Q9he23 neurospora
141	27	87.1	418	17	Q8TVF2	Q8tvf2 methanopyru	214	27	87.1	1079	5	Q17566	Q17566 caenorhabdi
142	27	87.1	424	16	Q9KRY6	Q9kry6 streptomyces	215	27	87.1	1094	5	Q23915	Q23915 dictyosteli
143	27	87.1	429	16	Q9A018	Q9a018 streptococc	216	27	87.1	1108	13	Q9PWP0	Q9pwp0 tetradon f
144	27	87.1	439	2	Q8RMF3	Q8rmf3 streptococc	217	27	87.1	1119	2	Q8ROE9	Q8roe9 thermus the
145	27	87.1	439	16	Q8XD63	Q8xd63 escherichia	218	27	87.1	1136	3	Q9HGK8	Q9hgk8 tuber borch
146	27	87.1	445	17	Q9HK59	Q9hk59 thermoplasm	219	27	87.1	1137	5	Q93250	Q93250 caenorhabdi
147	27	87.1	447	4	Q96AK9	Q96ak9 homo sapien	220	27	87.1	1157	3	Q9HE10	Q9he10 blumeria gr
148	27	87.1	453	16	Q93MB2	Q93mb2 clostridium	221	27	87.1	1265	4	Q9E267	Q9e267 homo sapien
149	27	87.1	455	17	Q8U0L4	Q8u0l4 pyrococcus	222	27	87.1	1352	17	Q9S837	Q9s837 pyrococcus
150	27	87.1	456	16	Q9F9K9	Q9f9k9 pisciricket	223	27	87.1	1359	10	Q9SMW6	Q9smw6 arabidopsis
151	27	87.1	465	16	Q8U7K8	Q8u7k8 agrobacteri	224	27	87.1	1405	10	Q9LH88	Q9lh88 arabidopsis
152	27	87.1	472	2	Q43941	Q43941 acinetobact	225	27	87.1	1678	13	Q8UW88	Q8uw88 cryzias lat
153	27	87.1	474	16	Q9KSM2	Q9ksm2 vibrio chol	226	27	87.1	1833	5	Q9YME7	Q9yme7 drosophila
154	27	87.1	476	16	Q9KSM2	Q9ksm2 vibrio chol	227	27	87.1	2173	5	Q9U622	Q9u622 drosophila
155	27	87.1	478	5	Q8SKX1	Q8skx1 drosophila	228	27	87.1	2257	5	Q9Y132	Q9y132 drosophila
156	27	87.1	488	2	Q87253	Q87253 lactococcus	229	27	87.1	3187	5	Q9BLV5	Q9blv5 leishmania
157	27	87.1	488	2	Q48617	Q48617 lactococcus	230	27	87.1	5171	4	Q8WVK9	Q8wvk9 homo sapien
158	27	87.1	488	16	Q9CHY0	Q9chy0 lactococcus	231	27	87.1	10223	2	Q54296	Q54296 streptomyces
159	27	87.1	499	16	Q44318	Q44318 anabaena sp	232	27	87.1	105	17	Q97U37	Q97u37 sulfobus
160	27	87.1	499	16	Q8YKN8	Q8ykn8 anabaena sp	233	27	87.1	117	2	Q914C9	Q914c9 gamma prote
161	27	87.1	500	16	Q9ANB5	Q9anb5 bradyrhizob	234	27	87.1	117	2	Q914C9	Q914c9 gamma prote
162	27	87.1	504	16	Q9ABF5	Q9abf5 caulobacter	235	27	87.1	121	2	P95818	P95818 streptococc

236	26	83.9	123	6	09GMY9	09GMY9 macaca fasc	309	26	83.9	662	5	062239	062239 caenorhabdi
237	26	83.9	123	9	09G006	09G006 bacterioplasm	310	26	83.9	667	2	08VUHL	08VUHL streptococc
238	26	83.9	133	17	097AV4	097AV4 thermoplasm	311	26	83.9	671	16	09KVH9	09KVH9 vibrio chol
239	26	83.9	136	11	09PIU9	09PIU9 campylobact	312	26	83.9	686	10	09C602	09C602 arabidopsis
240	26	83.9	137	16	089076	089076 mus muscicul	313	26	83.9	691	10	08VXZ2	08VXZ2 arabidopsis
241	26	83.9	141	4	08WU0	08WU0 homo sapien	314	26	83.9	695	5	096441	096441 caenorhabdi
242	26	83.9	141	11	09DAX5	09DAX5 mus muscicu	315	26	83.9	695	5	09U8D7	09U8D7 chymomyza a
243	26	83.9	147	17	082WP9	082WP9 pyrobacteryu	316	26	83.9	699	17	08TWT4	08TWT4 methanopyru
244	26	83.9	148	10	08V250	08V250 arabidopsis	317	26	83.9	700	5	08WS94	08WS94 caenorhabdi
245	26	83.9	148	10	094A24	094A24 arabidopsis	318	26	83.9	702	17	09HMZ6	09HMZ6 thermoplasm
246	26	83.9	149	6	09BGT5	09BGT5 macaca fasc	319	26	83.9	706	5	09NDS3	09NDS3 dictyostell
247	26	83.9	157	15	09K408	09K408 streptomyc	320	26	83.9	715	10	08S2D3	08S2D3 oryza sativ
248	26	83.9	158	15	041878	041878 human immun	321	26	83.9	721	16	09A4A9	09A4A9 caulobacter
249	26	83.9	176	2	09KWC9	09KWC9 lactobacill	322	26	83.9	727	4	08TCC9	08TCC9 homo sapien
250	26	83.9	201	16	093CRO	093CRO shigella bo	323	26	83.9	727	11	091VD9	091VD9 mus muscicu
251	26	83.9	203	17	097J77	097J77 clostridium	324	26	83.9	729	5	09N4Y8	09N4Y8 caenorhabdi
252	26	83.9	203	16	0972T6	0972T6 sulfobolus	325	26	83.9	739	5	09V6J2	09V6J2 drosophila
253	26	83.9	213	11	08VHH1	08VHH1 notomys ale	326	26	83.9	751	5	09V6M9	09V6M9 drosophila
254	26	83.9	214	2	08RR80	08RR80 acetobacter	327	26	83.9	757	5	09VYH2	09VYH2 drosophila
255	26	83.9	221	17	09HJ83	09HJ83 thermoplasm	328	26	83.9	760	10	023388	023388 arabidopsis
256	26	83.9	221	17	08TW5	08TW5 methanopyru	329	26	83.9	762	2	09LCC0	09LCC0 alcaligenes
257	26	83.9	229	12	0914G5	0914G5 sulfobolus	330	26	83.9	766	10	09ER95	09ER95 arabidopsis
258	26	83.9	242	10	09LSC8	09LSC8 arabidopsis	331	26	83.9	780	17	09HJCS	09HJCS pseudomonas
259	26	83.9	249	5	09VBS3	09VBS3 drosophila	332	26	83.9	799	16	09HJ30	09HJ30 pseudomonas
260	26	83.9	253	16	0970Z1	0970Z1 streptococc	333	26	83.9	830	16	09HJY6	09HJY6 caenorhabdi
261	26	83.9	255	4	09P1A0	09P1A0 homo sapien	334	26	83.9	839	4	09UP83	09UP83 homo sapien
262	26	83.9	259	12	08QZ08	08QZ08 rana tigrin	335	26	83.9	849	10	09SHG6	09SHG6 arabidopsis
263	26	83.9	261	16	P74291	P74291 synechocyst	336	26	83.9	886	13	042588	042588 xenopus lae
264	26	83.9	262	12	092547	092547 epizootic h	337	26	83.9	904	11	008971	008971 mus muscicu
265	26	83.9	275	10	09LDD8	09LDD8 oryza sativ	338	26	83.9	947	5	09BIA3	09BIA3 caenorhabdi
266	26	83.9	281	16	08Y397	08Y397 raietonia s	339	26	83.9	975	10	09LMB4	09LMB4 arabidopsis
267	26	83.9	284	5	023758	023758 clonorchis	340	26	83.9	975	16	08YK12	08YK12 anaabaena sp
268	26	83.9	284	5	017888	017888 caenorhabdi	341	26	83.9	990	2	09RPL8	09RPL8 streptococc
269	26	83.9	286	16	09AC04	09AC04 staphylococc	342	26	83.9	1054	10	09FKL3	09FKL3 arabidopsis
270	26	83.9	302	6	091416	091416 pseudomonas	343	26	83.9	1057	11	091X04	091X04 mus muscicu
271	26	83.9	303	6	095KC6	095KC6 macaca fasc	344	26	83.9	1139	10	08S665	08S665 oryza sativ
272	26	83.9	306	5	018585	018585 caenorhabdi	345	26	83.9	1226	16	08Y745	08Y745 bruceella me
273	26	83.9	308	17	08TN24	08TN24 methanosarc	346	26	83.9	1348	5	09VDT3	09VDT3 drosophila
274	26	83.9	315	4	096J26	096J26 homo sapien	347	26	83.9	1397	4	09NTG2	09NTG2 homo sapien
275	26	83.9	315	4	096S07	096S07 homo sapien	348	26	83.9	1450	10	09SJR6	09SJR6 arabidopsis
276	26	83.9	315	6	09EE05	09EE05 macaca fasc	349	26	83.9	1450	10	09SJR6	09SJR6 arabidopsis
277	26	83.9	315	6	095J75	095J75 macaca fasc	350	26	83.9	2464	5	09VEE5	09VEE5 drosophila
278	26	83.9	334	16	08RI75	08RI75 fusobacteri	351	26	83.9	2548	4	09UNJ2	09UNJ2 homo sapien
279	26	83.9	339	12	09YWT5	09YWT5 melanoplus	352	26	83.9	2626	11	09ZIN3	09ZIN3 rattus norv
280	26	83.9	344	16	097SC7	097SC7 streptococc	353	26	83.9	3988	17	08TPE1	08TPE1 methanosarc
281	26	83.9	346	5	09V6V9	09V6V9 drosophila	354	26	83.9	4099	10	09C7Z6	09C7Z6 arabidopsis
282	26	83.9	365	17	026986	026986 methanobact	355	26	83.9	16	8	09T206	09T206 solanum tub
283	26	83.9	368	2	0910S9	0910S9 rhizobium m	356	26	80.6	72	10	09LXJ2	09LXJ2 arabidopsis
284	26	83.9	378	16	08UC02	08UC02 agrobacteri	357	26	80.6	84	12	09DP94	09DP94 kunjin viru
285	26	83.9	382	2	095380	095380 escherichia	358	26	80.6	85	12	082986	082986 kunjin viru
286	26	83.9	382	2	08RLK9	08RLK9 salmonella	359	26	80.6	88	12	09DP22	09DP22 kunjin viru
287	26	83.9	384	16	09HMK9	09HMK9 pseudomonas	360	26	80.6	92	15	09DP21	09DP21 human immun
288	26	83.9	402	4	09H766	09H766 homo sapien	361	26	80.6	92	15	076999	076999 human immun
289	26	83.9	408	3	094693	094693 schizosacch	362	26	80.6	92	15	077000	077000 human immun
290	26	83.9	413	17	051984	051984 halobacteri	363	26	80.6	100	17	08TR13	08TR13 human immun
291	26	83.9	441	13	093403	093403 lorpedo cal	364	26	80.6	102	17	058485	058485 methanosarc
292	26	83.9	449	10	09LS16	09LS16 arabidopsis	365	26	80.6	104	2	09AMR2	09AMR2 pyrococcus
293	26	83.9	449	13	093346	093346 gallus gall	366	26	80.6	104	2	09AMR2	09AMR2 pyrococcus
294	26	83.9	458	17	059215	059215 pyrococcus	367	26	80.6	114	16	08YDL3	08YDL3 bruceella ab
295	26	83.9	470	5	09N8T1	09N8T1 typanosoma	368	26	80.6	117	9	038510	038510 bacterioplasm
296	26	83.9	476	13	093315	093315 cyprinlus ca	369	26	80.6	117	9	038511	038511 bacterioplasm
297	26	83.9	485	17	09HLM6	09HLM6 thermoplasm	370	26	80.6	117	9	038512	038512 bacterioplasm
298	26	83.9	488	10	09LNG1	09LNG1 arabidopsis	371	26	80.6	117	9	038513	038513 bacterioplasm
299	26	83.9	498	11	064596	064596 rattus norv	372	26	80.6	117	9	038514	038514 bacterioplasm
300	26	83.9	507	16	09POC4	09POC4 ureaplasma	373	26	80.6	117	9	038515	038515 bacterioplasm
301	26	83.9	542	4	03H8T5	03H8T5 homo sapien	374	26	80.6	117	9	038516	038516 bacterioplasm
302	26	83.9	579	12	08OMW8	08OMW8 cowpox viru	375	26	80.6	117	9	038517	038517 bacterioplasm
303	26	83.9	581	12	072760	072760 cowpox viru	376	26	80.6	122	10	09SVU5	09SVU5 arabidopsis
304	26	83.9	596	16	097HW2	097HW2 clostridium	377	26	80.6	126	17	09HK12	09HK12 human immun
305	26	83.9	615	16	09PCG7	09PCG7 xyliella fas	378	26	80.6	130	15	08UNP0	08UNP0 human immun
306	26	83.9	632	5	08SX70	08SX70 drosophila	379	26	80.6	133	16	09HMK6	09HMK6 pseudomonas
307	26	83.9	637	17	0972X6	0972X6 sulfobolus	380	26	80.6	141	2	044784	044784 borrelia bu
308	26	83.9	653	16	067097	067097 aquifex aeo	381	26	80.6	141	2	044784	044784 borrelia bu

382	25	80.6	141	2	Q9R3H3	Q9R3H3 borrella bu	455	25	80.6	372	16	Q67770	Q67770 aquifex aeo
383	25	80.6	141	2	Q9S0J3	Q9S0J3 borrella bu	456	25	80.6	374	16	Q8YRS1	Q8YRS1 anabaena sp
384	25	80.6	141	2	Q9S0A9	Q9S0A9 borrella bu	457	25	80.6	378	5	Q62012	Q62012 caenorhabdi
385	25	80.6	141	2	Q9S037	Q9S037 borrella bu	458	25	80.6	386	10	Q94776	Q94776 oryza sativ
386	25	80.6	141	2	Q9RXY5	Q9RXY5 borrella bu	459	25	80.6	388	13	Q90XK2	Q90XK2 brachydanio
387	25	80.6	141	2	Q07496	Q07496 borrella bu	460	25	80.6	390	5	Q9YPM4	Q9YPM4 drosophila
388	25	80.6	141	2	Q04788	Q04788 borrella bu	461	25	80.6	390	5	Q8SKR2	Q8SKR2 drosophila
389	25	80.6	141	2	Q04478	Q04478 borrella bu	462	25	80.6	392	3	Q9P883	Q9P883 aaracus di
390	25	80.6	141	2	Q9R9B8	Q9R9B8 borrella bu	463	25	80.6	393	2	Q9ROF1	Q9ROF1 mycoplasma
391	25	80.6	142	5	Q9VRC6	Q9VRC6 drosophila	464	25	80.6	394	16	Q97GV5	Q97GV5 clostridium
392	25	80.6	145	16	Q9H2T9	Q9H2T9 pseudomonas	465	25	80.6	397	5	Q8R0L0	Q8R0L0 drosophila
393	25	80.6	147	16	Q9X080	Q9X080 thermotoga	466	25	80.6	398	2	Q93P60	Q93P60 acholeplasm
394	25	80.6	162	16	Q94432	Q94432 ciona inte	467	25	80.6	401	11	Q91XC7	Q91XC7 mus muscu
395	25	80.6	162	16	Q03437	Q03437 bacillus su	468	25	80.6	402	16	Q8C106	Q8C106 lactococcus
396	25	80.6	163	16	Q92TW1	Q92TW1 rhizobium m	469	25	80.6	403	16	Q48345	Q48345 anabaena sp
397	25	80.6	183	16	Q8R5V3	Q8R5V3 thermoaer	470	25	80.6	402	16	Q97P00	Q97P00 streptococ
398	25	80.6	185	9	Q64022	Q64022 bacterioph	471	25	80.6	403	16	Q8YX05	Q8YX05 anabaena sp
399	25	80.6	185	16	Q03199	Q03199 bacillus su	472	25	80.6	403	16	Q8YNU6	Q8YNU6 anabaena sp
400	25	80.6	196	5	Q8SW34	Q8SW34 encephalit	473	25	80.6	403	16	Q8YKH3	Q8YKH3 anabaena sp
401	25	80.6	200	16	Q8UFX7	Q8UFX7 agrobacteri	474	25	80.6	410	11	Q9D5F1	Q9D5F1 mus muscu
402	25	80.6	203	16	Q8YBS4	Q8YBS4 bruceella me	475	25	80.6	413	16	Q8C177	Q8C177 pasteurella
403	25	80.6	208	4	Q75041	Q75041 homo sapien	476	25	80.6	413	16	Q8YLT9	Q8YLT9 anabaena sp
404	25	80.6	209	5	Q667H3	Q667H3 hydra atten	477	25	80.6	418	17	Q8YK44	Q8YK44 methanosarc
405	25	80.6	209	12	Q80L21	Q80L21 sulfolobus	478	25	80.6	418	17	Q96U29	Q96U29 glibberella
406	25	80.6	210	16	Q8XYV4	Q8XYV4 bruceella me	479	25	80.6	420	3	Q9J310	Q9J310 mus muscu
407	25	80.6	216	4	Q9SDH0	Q9SDH0 homo sapien	480	25	80.6	425	17	Q97W07	Q97W07 sulfolobus
408	25	80.6	222	12	Q9S0R6	Q9S0R6 myxoma viru	481	25	80.6	425	17	Q9S3P9	Q9S3P9 arabisdops
409	25	80.6	234	5	Q9X1W7	Q9X1W7 entodinium	482	25	80.6	436	10	Q9XEP9	Q9XEP9 lycopersico
410	25	80.6	235	16	Q928V7	Q928V7 listeria in	483	25	80.6	439	10	Q9XEP9	Q9XEP9 arabisdops
411	25	80.6	238	17	Q02857	Q02857 archaeoglob	484	25	80.6	441	10	Q9XE48	Q9XE48 lycopersico
412	25	80.6	241	16	Q97H70	Q97H70 clostridium	485	25	80.6	442	10	Q9XE50	Q9XE50 lycopersico
413	25	80.6	248	5	Q96945	Q96945 geodia cydo	486	25	80.6	442	10	P92961	P92961 arabisdops
414	25	80.6	248	11	Q9J120	Q9J120 mus muscu	487	25	80.6	442	16	Q9PGK3	Q9PGK3 xylella fas
415	25	80.6	254	16	Q87599	Q87599 streptomyc	488	25	80.6	444	5	Q9YXK2	Q9YXK2 caenorhabdi
416	25	80.6	258	16	Q9X1Z8	Q9X1Z8 thermotoga	489	25	80.6	450	10	Q94877	Q94877 hordium vul
417	25	80.6	261	17	Q9V3X3	Q9V3X3 drosophila	490	25	80.6	450	16	Q91S64	Q91S64 pseudomonas
418	25	80.6	261	17	Q9H1W9	Q9H1W9 thermoplas	491	25	80.6	453	10	Q91KRI	Q91KRI atriplex ho
419	25	80.6	262	10	Q940K6	Q940K6 arabisdops	492	25	80.6	456	5	Q45301	Q45301 caenorhabdi
420	25	80.6	268	16	Q9S1T3	Q9S1T3 streptomyc	493	25	80.6	458	5	Q8SR56	Q8SR56 encephalit
421	25	80.6	272	17	Q9H143	Q9H143 thermoplas	494	25	80.6	464	5	Q9U1R5	Q9U1R5 caenorhabdi
422	25	80.6	274	16	Q9S3H2	Q9S3H2 rhizobium l	495	25	80.6	465	11	Q9D0L8	Q9D0L8 mus muscu
423	25	80.6	276	4	Q96F02	Q96F02 homo sapien	496	25	80.6	467	5	Q9N8B5	Q9N8B5 trypanosoma
424	25	80.6	284	17	Q90202	Q90202 pyrococcus	497	25	80.6	467	16	Q9K1Z0	Q9K1Z0 neisseria m
425	25	80.6	286	5	Q23011	Q23011 caenorhabdi	498	25	80.6	469	11	Q9N9M5	Q9N9M5 mus muscu
426	25	80.6	286	17	Q9YAJ9	Q9YAJ9 aeropyrum p	499	25	80.6	469	5	Q8WPM7	Q8WPM7 paramecium
427	25	80.6	305	16	Q970Y0	Q970Y0 streptococ	500	25	80.6	471	16	Q8YBK0	Q8YBK0 bruceella me
428	25	80.6	306	16	Q9HY52	Q9HY52 pseudomonas	501	25	80.6	473	10	Q9LUN7	Q9LUN7 oryza sativ
429	25	80.6	313	2	Q9L8R2	Q9L8R2 pseudomonas	502	25	80.6	476	4	Q9UTU9	Q9UTU9 homo sapien
430	25	80.6	316	5	Q9XMH8	Q9XMH8 caenorhabdi	503	25	80.6	476	4	Q43148	Q43148 homo sapien
431	25	80.6	316	5	Q9VHH0	Q9VHH0 drosophila	504	25	80.6	481	10	Q81498	Q81498 arabisdops
432	25	80.6	318	16	Q9U061	Q9U061 staphylococ	505	25	80.6	482	16	Q8XU27	Q8XU27 raietonia s
433	25	80.6	326	15	Q9HTT2	Q9HTT2 pseudomonas	506	25	80.6	487	10	Q92RC8	Q92RC8 carica papa
434	25	80.6	328	5	Q9V4H8	Q9V4H8 drosophila	507	25	80.6	488	10	Q49819	Q49819 carica papa
435	25	80.6	331	5	Q9XU69	Q9XU69 caenorhabdi	508	25	80.6	491	5	Q9GV11	Q9GV11 ephydalia f
436	25	80.6	331	10	Q9SX81	Q9SX81 arabisdops	509	25	80.6	494	3	Q9C1E2	Q9C1E2 botrytis ci
437	25	80.6	332	2	Q69231	Q69231 bacillus sp	510	25	80.6	496	3	Q9H2R0	Q9H2R0 botrytis ci
438	25	80.6	333	10	Q9FFA1	Q9FFA1 arabisdops	511	25	80.6	498	3	Q9H2R0	Q9H2R0 botrytis ci
439	25	80.6	333	16	Q9F3K0	Q9F3K0 streptomyc	512	25	80.6	498	3	Q9H2R1	Q9H2R1 botrytis ci
440	25	80.6	340	10	Q9FL46	Q9FL46 arabisdops	513	25	80.6	498	3	Q9C1E3	Q9C1E3 botrytis ci
441	25	80.6	346	16	Q07414	Q07414 mycobacteri	514	25	80.6	498	16	Q9RTW9	Q9RTW9 delinococcus
442	25	80.6	348	4	Q96GZ7	Q96GZ7 homo sapien	515	25	80.6	500	3	Q9H2R5	Q9H2R5 botrytis ci
443	25	80.6	350	16	Q05343	Q05343 mycobacteri	516	25	80.6	500	3	Q9C1E5	Q9C1E5 botrytis ci
444	25	80.6	351	2	Q9RA92	Q9RA92 sinorhizobi	517	25	80.6	500	3	Q9H2R4	Q9H2R4 caenorhabdi
445	25	80.6	351	17	Q97W55	Q97W55 sulfolobus	518	25	80.6	501	3	Q9H2R4	Q9H2R4 botrytis ci
446	25	80.6	357	16	Q8UDH4	Q8UDH4 agrobacteri	519	25	80.6	501	3	Q9H2R3	Q9H2R3 botrytis ci
447	25	80.6	359	16	Q8R828	Q8R828 fusobacteri	520	25	80.6	501	3	Q9H2R3	Q9H2R3 botrytis ci
448	25	80.6	363	16	P73886	P73886 synochocyst	521	25	80.6	501	3	Q9C1E4	Q9C1E4 botrytis ci
449	25	80.6	365	16	P73756	P73756 synochocyst	522	25	80.6	502	3	Q9C1E4	Q9C1E4 botrytis ci
450	25	80.6	367	10	Q9M6B1	Q9M6B1 chlamydomon	523	25	80.6	503	3	Q9H2R2	Q9H2R2 botrytis ci
451	25	80.6	367	10	Q9M6B0	Q9M6B0 chlamydomon	524	25	80.6	504	3	Q9C1E6	Q9C1E6 botrytis ci
452	25	80.6	369	7	Q96W00	Q96W00 geotrichum	525	25	80.6	504	4	Q94996	Q94996 homo sapien
453	25	80.6	370	17	Q29775	Q29775 archaeoglob	526	25	80.6	505	5	Q9YVH4	Q9YVH4 drosophila
454	25	80.6	371	12	Q69294	Q69294 gallid hep	527	25	80.6				

528	25	80.6	509	5	Q95W10	Q95W10 drosophila	601	25	80.6	970	12	Q80Z06	Q80Z06 rana tigris
529	25	80.6	512	15	Q9WM14	Q9WM14 raltsonia s	602	25	80.6	971	16	Q9ZE42	Q9ZE42 rickettsia
530	25	80.6	512	16	Q8XYX0	Q8XYX0 turkey aden	603	25	80.6	976	16	Q97EV0	Q97EV0 clostridium
531	25	80.6	517	12	Q9YU51	Q9YU51 turkey aden	604	25	80.6	979	16	Q98PK8	Q98PK8 mycoplasma
532	25	80.6	522	3	Q93867	Q93867 erysiphe gr	605	25	80.6	983	10	Q9EP33	Q9EP33 oryza sativ
533	25	80.6	522	3	Q9P428	Q9P428 botrytis ci	606	25	80.6	1000	17	Q96YV3	Q96YV3 sulfolobus
534	25	80.6	522	3	Q8TGV1	Q8TGV1 monilia f	607	25	80.6	1058	5	Q96J35	Q96J35 drosophila
535	25	80.6	525	5	Q9VW60	Q9VW60 drosophila	608	25	80.6	1075	5	Q95U75	Q95U75 drosophila
536	25	80.6	534	5	Q21063	Q21063 caenorhabdi	609	25	80.6	1115	2	Q93R34	Q93R34 streptococ
537	25	80.6	535	10	Q9S127	Q9S127 arabadopsi	610	25	80.6	1118	4	Q9UIM0	Q9UIM0 homo sapien
538	25	80.6	536	17	Q97YB2	Q97YB2 sulfolobus	611	25	80.6	1133	10	Q82504	Q82504 arabadopsi
539	25	80.6	539	5	Q96Z08	Q96Z08 sulfolobus	612	25	80.6	1185	10	Q9AY94	Q9AY94 oryza sativ
540	25	80.6	539	5	Q9YKM4	Q9YKM4 drosophila	613	25	80.6	1205	5	Q96694	Q96694 plasmodium
541	25	80.6	542	2	Q96294	Q96294 thibacilli	614	25	80.6	1255	5	Q8SVZ5	Q8SVZ5 encephalito
542	25	80.6	548	2	Q86288	Q86288 lactococcus	615	25	80.6	1257	12	Q8QYV6	Q8QYV6 euprosterna
543	25	80.6	549	10	Q93Z39	Q93Z39 arabadopsi	616	25	80.6	1328	6	Q97836	Q97836 syncerus ca
544	25	80.6	550	16	Q25060	Q25060 helicobacte	617	25	80.6	1332	6	Q97897	Q97897 tragelaphus
545	25	80.6	550	16	Q92ME9	Q92ME9 helicobacte	618	25	80.6	1332	6	Q95325	Q95325 bos taurus
546	25	80.6	556	16	Q8XZ51	Q8XZ51 raltsonia s	619	25	80.6	1401	10	Q94LJ6	Q94LJ6 oryza sativ
547	25	80.6	558	5	Q19290	Q19290 caenorhabdi	620	25	80.6	1416	10	Q8S609	Q8S609 oryza sativ
548	25	80.6	568	5	Q9V903	Q9V903 drosophila	621	25	80.6	1450	10	Q9LPH0	Q9LPH0 xenopus lae
549	25	80.6	575	16	Q8YMW7	Q8YMW7 arabidopsi	622	25	80.6	1490	13	P79322	P79322 xenopus lae
550	25	80.6	577	5	Q9HFR9	Q9HFR9 escherichia	623	25	80.6	1497	10	Q8S718	Q8S718 oryza sativ
551	25	80.6	584	5	Q25283	Q25283 leishmania	624	25	80.6	1529	3	Q9Y7C6	Q9Y7C6 aspergillus
552	25	80.6	588	16	Q99UV8	Q99UV8 staphylococ	625	25	80.6	1676	10	Q23332	Q23332 arabadopsi
553	25	80.6	589	5	Q19289	Q19289 caenorhabdi	626	25	80.6	1680	10	Q64588	Q64588 arabadopsi
554	25	80.6	594	4	Q96HN9	Q96HN9 homo sapien	627	25	80.6	1720	10	Q9AY68	Q9AY68 oryza sativ
555	25	80.6	595	10	Q9XEX0	Q9XEX0 nicotiana t	628	25	80.6	1769	5	Q16625	Q16625 caenorhabdi
556	25	80.6	595	16	Q9A9Z2	Q9A9Z2 caulobacter	629	25	80.6	2183	12	Q98333	Q98333 measles vir
557	25	80.6	603	10	Q9SZT9	Q9SZT9 arabadopsi	630	25	80.6	2233	12	Q9U7D6	Q9U7D6 bovine para
558	25	80.6	603	11	Q9UUC5	Q9UUC5 mus musculu	631	25	80.6	2233	12	Q9U7D0	Q9U7D0 bovine para
559	25	80.6	607	12	Q8Q5A2	Q8Q5A2 felis domes	632	25	80.6	2233	12	Q81890	Q81890 human para
560	25	80.6	609	5	Q962H9	Q962H9 toxoplasma	633	25	80.6	2258	12	Q81081	Q81081 human para
561	25	80.6	614	10	Q94DX2	Q94DX2 oryza sativ	634	25	80.6	2556	12	Q9CH56	Q9CH56 gallid hep
562	25	80.6	615	5	Q961X3	Q961X3 drosophila	635	25	80.6	7525	2	Q9K1E0	Q9K1E0 streptomyce
563	25	80.6	617	17	Q9HKM1	Q9HKM1 thermoplas	636	25	80.6	7576	2	Q9ZG44	Q9ZG44 streptomyce
564	25	80.6	619	5	Q8ST82	Q8ST82 encephalito	637	25	77.4	42	4	Q9WMW6	Q9WMW6 japanese en
565	25	80.6	632	10	Q8W3I6	Q8W3I6 oryza sativ	638	25	77.4	38	4	Q9H453	Q9H453 homo sapien
566	25	80.6	633	16	Q9VZ84	Q9VZ84 drosophila	639	25	77.4	53	17	Q9YCT4	Q9YCT4 aeropyrum p
567	25	80.6	643	15	Q8XKX0	Q8XKX0 clostridium	640	25	77.4	55	5	Q9GNZ8	Q9GNZ8 plasmodium
568	25	80.6	654	10	Q8S7H9	Q8S7H9 oryza sativ	641	25	77.4	55	5	Q9GZ89	Q9GZ89 plasmodium
569	25	80.6	664	10	Q9STY9	Q9STY9 arabadopsi	642	25	77.4	55	5	Q9GZ83	Q9GZ83 plasmodium
570	25	80.6	669	17	Q97CU9	Q97CU9 thermoplas	643	25	77.4	55	12	Q9NMW7	Q9NMW7 japanese en
571	25	80.6	671	5	Q9MYV6	Q9MYV6 penaeus van	644	25	77.4	58	5	Q9GNE0	Q9GNE0 plasmodium
572	25	80.6	678	12	Q8V3A7	Q8V3A7 caspalia e	645	25	77.4	58	5	Q9GNB9	Q9GNB9 plasmodium
573	25	80.6	678	17	Q8TMR7	Q8TMR7 methanosarc	646	25	77.4	58	5	Q9GN91	Q9GN91 plasmodium
574	25	80.6	689	16	Q9KY46	Q9KY46 streptomyce	647	25	77.4	58	5	Q9GN77	Q9GN77 plasmodium
575	25	80.6	696	5	Q8SR54	Q8SR54 encephalito	648	25	77.4	58	5	Q9GN76	Q9GN76 plasmodium
576	25	80.6	698	11	Q8R3J2	Q8R3J2 mus musculu	649	25	77.4	58	5	Q9GZB8	Q9GZB8 plasmodium
577	25	80.6	699	5	Q95ZC2	Q95ZC2 leishmania	650	25	77.4	58	5	Q9GZB7	Q9GZB7 plasmodium
578	25	80.6	699	11	Q9DA46	Q9DA46 mus musculu	651	25	77.4	58	5	Q9GZB6	Q9GZB6 plasmodium
579	25	80.6	730	12	Q91TML	Q91TML tupala hep	652	25	77.4	58	5	Q9GZB2	Q9GZB2 plasmodium
580	25	80.6	735	10	Q93Z03	Q93Z03 arabadopsi	653	25	77.4	58	5	Q9GZB1	Q9GZB1 plasmodium
581	25	80.6	736	10	Q82754	Q82754 arabadopsi	654	25	77.4	58	5	Q9GZA9	Q9GZA9 plasmodium
582	25	80.6	754	4	Q9P0L1	Q9P0L1 homo sapien	655	25	77.4	58	5	Q9GZAS	Q9GZAS plasmodium
583	25	80.6	764	17	Q8TPU1	Q8TPU1 methanosarc	656	25	77.4	58	5	Q9GZAS	Q9GZAS plasmodium
584	25	80.6	778	5	Q9BLD4	Q9BLD4 palinopecte	657	25	77.4	58	5	Q9GZAS	Q9GZAS plasmodium
585	25	80.6	785	10	Q9FRN0	Q9FRN0 oryza sativ	658	25	77.4	58	5	Q9GZAS	Q9GZAS plasmodium
586	25	80.6	786	16	Q9FBS2	Q9FBS2 streptomyce	659	25	77.4	58	5	Q9GZ97	Q9GZ97 plasmodium
587	25	80.6	787	17	Q9UYH2	Q9UYH2 pyrococcus	660	25	77.4	58	5	Q9GZ96	Q9GZ96 plasmodium
588	25	80.6	801	16	Q9YX11	Q9YX11 anabaena sp	661	25	77.4	58	5	Q9GZ94	Q9GZ94 plasmodium
589	25	80.6	806	5	Q8T2A3	Q8T2A3 dictyosteli	662	25	77.4	58	5	Q9GZ93	Q9GZ93 plasmodium
590	25	80.6	832	5	Q963E6	Q963E6 drosophila	663	25	77.4	58	5	Q9GZ91	Q9GZ91 plasmodium
591	25	80.6	847	16	Q92KG2	Q92KG2 rhizobium m	664	25	77.4	58	5	Q9GZ90	Q9GZ90 plasmodium
592	25	80.6	848	11	Q99P14	Q99P14 mus musculu	665	25	77.4	58	5	Q9GZ87	Q9GZ87 plasmodium
593	25	80.6	866	16	Q8UC06	Q8UC06 agrobacteri	666	25	77.4	58	5	Q9GZ84	Q9GZ84 plasmodium
594	25	80.6	869	2	Q9R912	Q9R912 rhizobium m	667	25	77.4	58	5	Q9GZ82	Q9GZ82 plasmodium
595	25	80.6	884	9	Q9T145	Q9T145 bacterioph	668	25	77.4	58	5	Q9GZ81	Q9GZ81 plasmodium
596	25	80.6	897	5	Q9XZRO	Q9XZRO caenorhabdi	669	25	77.4	58	5	Q9GZ80	Q9GZ80 plasmodium
597	25	80.6	901	5	Q9TX17	Q9TX17 caenorhabdi	670	25	77.4	58	5	Q9GZ78	Q9GZ78 plasmodium
598	25	80.6	929	16	Q86810	Q86810 streptomyce	671	25	77.4	58	5	Q9GZ77	Q9GZ77 plasmodium
599	25	80.6	938	5	Q9V8V8	Q9V8V8 drosophila	672	25	77.4	64	12	Q72732	Q72732 cowpox viru
600	25	80.6	954	5	Q9U0M5	Q9U0M5 plasmodium	673	25	77.4	65	12	Q82900	Q82900 japanese en

674	24	77.4	66	5	Q9M1P6	Q9M1P6 drosophila	747	24	77.4	164	16	Q9PNA4	Q9PNA4 campylobact
675	24	77.4	68	12	Q82899	Q82899 japanese en	748	24	77.4	165	16	Q91204	Q91204 pseudomonas
676	24	77.4	69	17	Q8TQC7	Q8TQC7 methanosarc	749	24	77.4	165	17	Q98810	Q98810 arthaeoglob
677	24	77.4	82	10	Q9L146	Q9L146 phaseolus v	750	24	77.4	166	10	Q04125	Q04125 nicotiana t
678	24	77.4	84	12	Q82898	Q82898 japanese en	751	24	77.4	167	2	Q8V1N8	Q8V1N8 pseudomonas
679	24	77.4	86	10	Q24439	Q24439 phaseolus v	752	24	77.4	169	5	Q20131	Q20131 caenorhabdi
680	24	77.4	88	12	Q9M1P6	Q9M1P6 japanese en	753	24	77.4	170	2	Q46103	Q46103 campylobact
681	24	77.4	89	16	Q9KUT7	Q9KUT7 vibrio chol	754	24	77.4	171	10	Q9FGY5	Q9FGY5 arabisdopsis
682	24	77.4	90	12	Q9M1W7	Q9M1W7 japanese en	755	24	77.4	171	10	Q92SW5	Q92SW5 hamamelis v
683	24	77.4	91	16	Q9K1V0	Q9K1V0 vibrio chol	756	24	77.4	172	10	Q46C8	Q46C8 theobroma c
684	24	77.4	94	2	Q32635	Q32635 helicobacte	757	24	77.4	173	10	Q46A1	Q46A1 theobroma c
685	24	77.4	95	10	Q9S7N8	Q9S7N8 glycine max	758	24	77.4	174	10	Q945X7	Q945X7 theobroma c
686	24	77.4	100	16	Q931X7	Q931X7 staphylococ	759	24	77.4	174	17	Q9V161	Q9V161 pyrococcus
687	24	77.4	102	13	Q91662	Q91662 xenopus lae	760	24	77.4	178	17	Q975A2	Q975A2 sulfolobus
688	24	77.4	110	16	Q8X5M1	Q8X5M1 escherichia	761	24	77.4	180	10	Q9STB9	Q9STB9 gracillariop
689	24	77.4	111	12	Q9DM7	Q9DM7 sporoptera	762	24	77.4	183	10	Q9SHU8	Q9SHU8 arabidopsis
690	24	77.4	111	17	Q976V1	Q976V1 sulfolobus	763	24	77.4	184	12	Q2YVS0	Q2YVS0 melanoplus
691	24	77.4	112	2	Q9L484	Q9L484 uncultured	764	24	77.4	185	16	Q67590	Q67590 aquifex aeo
692	24	77.4	112	2	Q9L483	Q9L483 uncultured	765	24	77.4	187	4	Q96Q77	Q96Q77 homo sapien
693	24	77.4	112	2	Q9L480	Q9L480 uncultured	766	24	77.4	189	8	Q9TCA5	Q9TCA5 nephroselm
694	24	77.4	112	2	Q9L478	Q9L478 uncultured	767	24	77.4	190	8	Q9MD76	Q9MD76 lupinus lut
695	24	77.4	112	2	Q9L477	Q9L477 uncultured	768	24	77.4	190	8	Q9MD58	Q9MD58 lupinus ang
696	24	77.4	114	5	Q22333	Q22333 caenorhabdi	769	24	77.4	190	8	Q9MFG5	Q9MFG5 lupinus ang
697	24	77.4	114	16	Q8X3N5	Q8X3N5 escherichia	770	24	77.4	190	8	Q9MFG7	Q9MFG7 lupinus mut
698	24	77.4	115	16	Q53334	Q53334 mycobacteri	771	24	77.4	190	8	Q9MFG6	Q9MFG6 lupinus alb
699	24	77.4	117	2	Q9L4C2	Q9L4C2 uncultured	772	24	77.4	190	16	Q9M1W4	Q9M1W4 pseudomonas
700	24	77.4	117	2	Q9L4C1	Q9L4C1 uncultured	773	24	77.4	192	8	Q9MFE5	Q9MFE5 beta vulgar
701	24	77.4	117	2	Q9L4C0	Q9L4C0 uncultured	774	24	77.4	192	12	Q9Q3W5	Q9Q3W5 hepatitis c
702	24	77.4	117	2	Q9L4B9	Q9L4B9 uncultured	775	24	77.4	193	16	Q8XE96	Q8XE96 escherichia
703	24	77.4	117	2	Q9L4B8	Q9L4B8 uncultured	776	24	77.4	193	17	Q97U03	Q97U03 sulfolobus
704	24	77.4	117	2	Q9L4B6	Q9L4B6 uncultured	777	24	77.4	194	10	Q9L213	Q9L213 arabidopsis
705	24	77.4	117	2	Q9L4B5	Q9L4B5 uncultured	778	24	77.4	197	16	Q9A6X2	Q9A6X2 caulobacter
706	24	77.4	117	2	Q9L4B4	Q9L4B4 uncultured	779	24	77.4	199	16	P73231	P73231 synechocyst
707	24	77.4	117	2	Q9L4B3	Q9L4B3 uncultured	780	24	77.4	199	17	Q8U385	Q8U385 pyrococcus
708	24	77.4	117	2	Q9L4B1	Q9L4B1 uncultured	781	24	77.4	200	8	Q98RNO	Q98RNO guillardia
709	24	77.4	117	2	Q9L4A8	Q9L4A8 uncultured	782	24	77.4	200	15	Q992U2	Q992U2 streptococ
710	24	77.4	117	2	Q9L4A7	Q9L4A7 uncultured	783	24	77.4	200	15	Q8U862	Q8U862 agrobacteri
711	24	77.4	117	2	Q9L4A6	Q9L4A6 uncultured	784	24	77.4	201	15	Q67778	Q67778 aquifex aeo
712	24	77.4	117	2	Q9L4A2	Q9L4A2 uncultured	785	24	77.4	201	16	Q98KQ7	Q98KQ7 rhizobium l
713	24	77.4	117	2	Q9L4A1	Q9L4A1 uncultured	786	24	77.4	203	16	Q34487	Q34487 bacillus su
714	24	77.4	117	2	Q9L4A0	Q9L4A0 uncultured	787	24	77.4	204	5	Q26014	Q26014 plasmodium
715	24	77.4	117	2	Q9L4A0	Q9L4A0 uncultured	788	24	77.4	206	8	Q954U3	Q954U3 beta vulgar
716	24	77.4	117	2	Q9L4A9	Q9L4A9 uncultured	789	24	77.4	208	17	Q9UXD8	Q9UXD8 sulfolobus
717	24	77.4	117	2	Q9L4A8	Q9L4A8 uncultured	790	24	77.4	210	16	Q97FK9	Q97FK9 clostridium
718	24	77.4	117	5	Q9V1X6	Q9V1X6 drosophila	791	24	77.4	215	2	Q9XDT3	Q9XDT3 clostridium
719	24	77.4	117	17	Q97UX4	Q97UX4 sulfolobus	792	24	77.4	215	16	Q8X6R8	Q8X6R8 escherichia
720	24	77.4	121	2	Q48516	Q48516 leptospira	793	24	77.4	216	16	Q85823	Q85823 yersinia pe
721	24	77.4	125	4	Q12825	Q12825 homo sapien	794	24	77.4	216	17	Q8ZWP8	Q8ZWP8 pyrobaculum
722	24	77.4	126	16	Q8UD18	Q8UD18 agrobacteri	795	24	77.4	218	16	Q8ZD58	Q8ZD58 yersinia pe
723	24	77.4	127	16	Q8YR79	Q8YR79 anabaena sp	796	24	77.4	219	17	Q8U470	Q8U470 pyrococcus
724	24	77.4	127	16	Q9ZNG7	Q9ZNG7 staphylococ	797	24	77.4	220	16	Q9C1M2	Q9C1M2 pasteurella
725	24	77.4	130	4	Q14942	Q14942 homo sapien	798	24	77.4	222	17	Q975F4	Q975F4 sulfolobus
726	24	77.4	130	17	Q9YFC0	Q9YFC0 aeropyrum p	799	24	77.4	223	17	Q8TVX5	Q8TVX5 methanopyru
727	24	77.4	131	16	Q92CE2	Q92CE2 listeria in	800	24	77.4	225	16	Q8X1M6	Q8X1M6 clostridium
728	24	77.4	132	17	Q9H2K1	Q9H2K1 halobacteri	801	24	77.4	226	13	Q9PW60	Q9PW60 cyprinus ca
729	24	77.4	134	16	Q8R7K1	Q8R7K1 thermomaer	802	24	77.4	227	16	Q9EF23	Q9EF23 streptomyce
730	24	77.4	137	10	Q9L1P4	Q9L1P4 arabidopsis	803	24	77.4	229	10	Q43363	Q43363 nicotiana t
731	24	77.4	137	12	Q9P1M1	Q9P1M1 hepatitis c	804	24	77.4	230	2	Q931F2	Q931F2 propionibac
732	24	77.4	139	2	Q50234	Q50234 zymomonas m	805	24	77.4	230	13	Q8QGE7	Q8QGE7 cutunrix co
733	24	77.4	141	12	Q9Y1W6	Q9Y1W6 melanoplus	806	24	77.4	231	5	Q9U8G2	Q9U8G2 trypanosoma
734	24	77.4	142	16	Q9C1Y8	Q9C1Y8 iactococcus	807	24	77.4	231	16	Q53480	Q53480 mycobacteri
735	24	77.4	145	12	Q99ES0	Q99ES0 japanese en	808	24	77.4	232	16	Q821D1	Q821D1 yersinia pe
736	24	77.4	147	13	Q12940	Q12940 gallus gall	809	24	77.4	233	16	Q8Y4V0	Q8Y4V0 listeria mo
737	24	77.4	150	2	P74980	P74980 yersinia ps	810	24	77.4	233	16	Q8YCB2	Q8YCB2 brucella me
738	24	77.4	150	3	Q14087	Q14087 schizosacch	811	24	77.4	235	5	Q9NEM6	Q9NEM6 caenorhabdi
739	24	77.4	151	16	Q9ZE1P	Q9ZE1P streptomyce	812	24	77.4	236	4	Q9NMG0	Q9NMG0 homo sapien
740	24	77.4	151	10	Q99R14	Q99R14 staphylococ	813	24	77.4	237	16	Q8Y1W8	Q8Y1W8 anabaena sp
741	24	77.4	152	10	Q01920	Q01920 phytophthor	814	24	77.4	238	5	Q9B1Y9	Q9B1Y9 renilla mu
742	24	77.4	152	10	Q01918	Q01918 phytophthor	815	24	77.4	238	16	Q67863	Q67863 aquifex aeo
743	24	77.4	156	16	Q8R1P1	Q8R1P1 thermomaer	816	24	77.4	241	16	Q9KB70	Q9KB70 bacillus ha
744	24	77.4	157	5	Q9UTP2	Q9UTP2 pecten maxi	817	24	77.4	242	16	Q24951	Q24951 helicobacte
745	24	77.4	158	10	Q9A3L7	Q9A3L7 caulobacter	818	24	77.4	243	16	Q24951	Q24951 helicobacte
746	24	77.4	160	10	Q9MAW2	Q9MAW2 bruguliera g	819	24	77.4	243	3	Q12134	Q12134 saccharomyc

820	24	77.4	243	9	064333	064333 bacterioph	893	24	77.4	297	16	050726	050726 borrelia bu
821	24	77.4	244	16	08Y3V9	08Y3V9 listeria mo	894	24	77.4	297	16	050805	050805 borrelia bu
822	24	77.4	248	10	022210	022210 aradiidopsis	895	24	77.4	300	13	09P84	09P84 carassius a
823	24	77.4	248	16	09KCS1	09KCS1 bacillus ha	896	24	77.4	300	16	09W426	09W426 streptomyc
824	24	77.4	248	16	097DB9	097DB9 clostridium	897	24	77.4	301	10	09F539	09F539 zea mays (m
825	24	77.4	249	16	09K710	09K710 bacillus ha	898	24	77.4	301	16	09KX28	09KX28 vibrio chol
826	24	77.4	251	5	09NKM6	09NKM6 leishmania	899	24	77.4	302	10	041777	041777 zea mays (m
827	24	77.4	251	17	029354	029354 archaeoglob	900	24	77.4	303	4	09H317	09H317 homo sapien
828	24	77.4	252	10	040701	040701 oryza sativ	901	24	77.4	303	5	017706	017706 caenorhabd
829	24	77.4	253	16	08YV4	08YV4 escherichia	902	24	77.4	304	5	08S0A7	08S0A7 encephalito
830	24	77.4	254	10	065103	065103 oryza sativ	903	24	77.4	304	10	041776	041776 zea mays (m
831	24	77.4	254	11	09CT99	09CT99 mus musculu	904	24	77.4	304	10	08S5W8	08S5W8 oryza sativ
832	24	77.4	254	17	097M63	097M63 sulfolobus	905	24	77.4	309	2	09FDP8	09FDP8 staphylococ
833	24	77.4	255	16	09C118	09C118 lactococcus	906	24	77.4	309	16	024916	024916 helicobacte
834	24	77.4	255	16	08ZB31	08ZB31 yersinia pe	907	24	77.4	309	16	09ZMY1	09ZMY1 helicobacte
835	24	77.4	255	16	08RBP5	08RBP5 thermoaer	908	24	77.4	310	17	08U307	08U307 pyrococcus
836	24	77.4	258	2	066280	066280 escherichia	909	24	77.4	312	2	09L796	09L796 portphyromon
837	24	77.4	258	5	09GQ17	09GQ17 caenorhabd	910	24	77.4	312	4	095286	095286 homo sapien
838	24	77.4	259	16	09RXU2	09RXU2 deinococcus	911	24	77.4	312	16	098QY0	098QY0 mycoplasma
839	24	77.4	260	5	017343	017343 caenorhabd	912	24	77.4	315	2	09XWJ3	09XWJ3 pseudomonas
840	24	77.4	264	4	096AS3	096AS3 homo sapien	913	24	77.4	316	16	09S2H1	09S2H1 streptomyc
841	24	77.4	265	2	09KW39	09KW39 wolbachia's	914	24	77.4	319	2	093P84	093P84 comamonas t
842	24	77.4	265	10	004126	004126 nicotiana t	915	24	77.4	319	16	067018	067018 aquifex aeo
843	24	77.4	268	10	040457	040457 streptococ	916	24	77.4	319	16	097U08	097U08 sulfolobus
844	24	77.4	269	2	082875	082875 streptococ	917	24	77.4	320	16	09A7Z8	09A7Z8 callobacter
845	24	77.4	269	3	002392	002392 saccharomyc	918	24	77.4	321	11	09CXF2	09CXF2 mus musculu
846	24	77.4	269	5	018670	018670 drosophila	919	24	77.4	321	11	09CXF2	09CXF2 mus musculu
847	24	77.4	269	5	016076	016076 drosophila	920	24	77.4	322	12	09JUP8	09JUP8 vaccinia vl
848	24	77.4	269	5	016087	016087 drosophila	921	24	77.4	323	12	087633	087633 campox vitu
849	24	77.4	269	10	09SKW3	09SKW3 aradiidopsis	922	24	77.4	323	12	08Q072	08Q072 campox vitu
850	24	77.4	270	16	09RE14	09RE14 fusobacteri	923	24	77.4	323	12	08Q013	08Q013 campox vitu
851	24	77.4	271	2	047308	047308 escherichia	924	24	77.4	326	2	093653	093653 serratia ma
852	24	77.4	271	2	09K3D9	09K3D9 uncultured	925	24	77.4	327	5	09U2B5	09U2B5 caenorhabd
853	24	77.4	271	2	054477	054477 escherichia	926	24	77.4	327	16	069870	069870 streptomyc
854	24	77.4	271	10	005786	005786 klebsiella	927	24	77.4	328	16	08X115	08X115 salmonella
855	24	77.4	273	10	09STER	09STER aradiidopsis	928	24	77.4	329	2	09F6F5	09F6F5 bacteroides
856	24	77.4	273	16	08XU19	08XU19 clostridium	929	24	77.4	330	5	09NAP9	09NAP9 caenorhabd
857	24	77.4	274	16	09RNV2	09RNV2 pseudomonas	930	24	77.4	331	10	09FX56	09FX56 aradiidopsis
858	24	77.4	275	10	09MSA7	09MSA7 zea mays (m	931	24	77.4	331	10	09AW15	09AW15 guillardia
859	24	77.4	275	16	08RBU0	08RBU0 thermoaer	932	24	77.4	332	2	030670	030670 bacillus su
860	24	77.4	277	16	09CF12	09CF12 lactococcus	933	24	77.4	333	5	09BJ52	09BJ52 heterodera
861	24	77.4	278	10	041851	041851 zea mays (m	934	24	77.4	334	16	08YRP9	08YRP9 anapanea sp
862	24	77.4	279	16	053632	053632 synechocyst	935	24	77.4	335	5	09BLJ5	09BLJ5 acrythosiph
863	24	77.4	279	16	08XWS9	08XWS9 clostridium	936	24	77.4	335	16	09RUP4	09RUP4 chlamydia m
864	24	77.4	281	5	09U7A8	09U7A8 brachionus	937	24	77.4	336	5	09USE5	09USE5 entamoeba d
865	24	77.4	281	10	08RWD7	08RWD7 aradiidopsis	938	24	77.4	336	16	09BP68	09BP68 rhizobium l
866	24	77.4	281	17	0979W3	0979W3 thermoplasm	939	24	77.4	336	16	09W2Z8	09W2Z8 thermotoga
867	24	77.4	282	16	09KU52	09KU52 vibrio chol	940	24	77.4	339	17	058916	058916 pyrococcus
868	24	77.4	283	6	09BGG60	09BGG60 mustela put	941	24	77.4	339	17	08RH13	08RH13 methanosarc
869	24	77.4	284	17	09HUR5	09HUR5 thermoplasm	942	24	77.4	340	12	09YU06	09YU06 melonoplus
870	24	77.4	285	5	026013	026013 plasmodium	943	24	77.4	342	5	09SV86	09SV86 nosema locu
871	24	77.4	287	8	035980	035980 triticum ae	944	24	77.4	343	12	08Y543	08Y543 monkeypox v
872	24	77.4	290	4	09BSRO	09BSRO homo sapien	945	24	77.4	343	16	08ZP83	08ZP83 salmonella
873	24	77.4	290	5	09U7B4	09U7B4 adineta vag	946	24	77.4	343	16	08Z789	08Z789 salmonella
874	24	77.4	290	5	09NBE3	09NBE3 adineta vag	947	24	77.4	345	17	09HNW8	09HNW8 halobacteri
875	24	77.4	290	5	09NBE2	09NBE2 adineta vag	948	24	77.4	346	5	016328	016328 caenorhabd
876	24	77.4	291	5	09U7B3	09U7B3 habroclitrocha	949	24	77.4	347	16	099WX3	099WX3 staphylococ
877	24	77.4	291	5	09U7A5	09U7A5 philodina r	950	24	77.4	348	12	089110	089110 variola vir
878	24	77.4	291	5	09NBE6	09NBE6 esophora eh	951	24	77.4	348	12	089754	089754 variola vir
879	24	77.4	291	5	09NBE5	09NBE5 philodina r	952	24	77.4	348	12	089088	089088 variola vir
880	24	77.4	291	5	09NBE4	09NBE4 habroclitrocha	953	24	77.4	349	16	09ZEF4	09ZEF4 listeria in
881	24	77.4	291	5	09NBE4	09NBE4 habroclitrocha	954	24	77.4	349	12	08VX27	08VX27 campelpox vl
882	24	77.4	291	16	092937	092937 synechocyst	955	24	77.4	350	10	09CC9C	09CC9C aradiidopsis
883	24	77.4	291	12	097288	097288 sulfolobus	956	24	77.4	350	10	0941W2	0941W2 oryza sativ
884	24	77.4	292	12	036377	036377 alcelaphine	957	24	77.4	351	2	09ADY7	09ADY7 agrobacteri
885	24	77.4	292	12	055748	055748 chilo iride	958	24	77.4	351	10	09FKD1	09FKD1 aradiidopsis
886	24	77.4	293	5	09NBE1	09NBE1 philodina r	959	24	77.4	351	10	09EXS5	09EXS5 aradiidopsis
887	24	77.4	293	5	09NBE9	09NBE9 philodina r	960	24	77.4	352	4	096U72	096U72 homo sapien
888	24	77.4	295	16	09RTR9	09RTR9 deinococcus	961	24	77.4	352	4	059329	059329 comamonas t
889	24	77.4	296	5	09U7A7	09U7A7 brachionus	962	24	77.4	354	12	089406	089406 paramonacium
890	24	77.4	296	5	09U7A6	09U7A6 seison neba	963	24	77.4	354	12	005366	005366 actinobacil
891	24	77.4	296	17	027257	027257 methanobact	964	24	77.4	355	2	066249	066249 actinobacil
892	24	77.4	297	4	09F0R4	09F0R4 homo sapien	965	24	77.4	355	2	066249	066249 actinobacil

966	24	77.4	355	2	Q9XDO8	Q9XDO8 actinobacill
967	24	77.4	355	2	Q9JRS	Q9JRS actinobacill
968	24	77.4	355	2	Q9AQC6	Q9AQC6 actinobacill
969	24	77.4	355	16	Q8UCD6	Q8UCD6 agrobacteri
970	24	77.4	356	16	Q99WH9	Q99WH9 staphylococ
971	24	77.4	356	16	Q99WH5	Q99WH5 staphylococ
972	24	77.4	356	5	Q9NRS5	Q9NRS5 drosophila
973	24	77.4	359	5	Q9NRS3	Q9NRS3 drosophila
974	24	77.4	359	5	Q9NRS1	Q9NRS1 drosophila
975	24	77.4	360	5	Q20723	Q20723 caenorhabdi
976	24	77.4	361	2	Q85776	Q85776 rhizobium l
977	24	77.4	361	16	Q9KBP5	Q9KBP5 bacillus ha
978	24	77.4	362	2	Q9S162	Q9S162 erysipelothe
979	24	77.4	362	2	Q9S160	Q9S160 erysipelothe
980	24	77.4	362	8	Q32831	Q32831 pyrammonas
981	24	77.4	365	2	Q9R6W8	Q9R6W8 anabaena sp
982	24	77.4	365	5	Q9XTE4	Q9XTE4 caenorhabdi
983	24	77.4	366	2	Q9X406	Q9X406 methylosulf
984	24	77.4	366	16	Q51031	Q51031 borrelia bu
985	24	77.4	367	5	Q17575	Q17575 caenorhabdi
986	24	77.4	368	17	Q8ZVP6	Q8ZVP6 pyrobaculum
987	24	77.4	370	5	Q8SRH4	Q8SRH4 encephalito
988	24	77.4	370	16	Q66865	Q66865 aquifex ae
989	24	77.4	372	4	Q15391	Q15391 homo sapien
990	24	77.4	372	4	Q96CS5	Q96CS5 homo sapien
991	24	77.4	373	17	Q8U163	Q8U163 pyrococcus
992	24	77.4	375	5	Q21903	Q21903 caenorhabdi
993	24	77.4	375	10	Q91TF5	Q91TF5 arabidopsis
994	24	77.4	375	17	Q8RTZ0	Q8RTZ0 pyrococcus
995	24	77.4	376	4	Q96S06	Q96S06 homo sapien
996	24	77.4	377	16	Q8YF2	Q8YF2 anabaena sp
997	24	77.4	378	5	Q9NDU9	Q9NDU9 caenorhabdi
998	24	77.4	379	2	Q48738	Q48738 lactococcus
999	24	77.4	380	16	Q9CK63	Q9CK63 pasteurella
1000	24	77.4	380	16	Q8R69	Q8R69 thermomanaer

ALIGNMENTS

RESULT 1

Q9DOK5 PRELIMINARY; PRT; 114 AA.

AC Q9DOK5; 01-JUN-2001 (TREMBlrel. 17, Created)

DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE 2610008E20R1X protein.

GN CHER1 OR 2610008E20R1X.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC SURFAC=57BL/6J; TISSUE=EMBRYO;

RA MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshio M., Itoh M., Ishi Y.,

RA Sakai K., Okada T., Furuno M., Aono H., Baldarelli R., Barsch G.,

RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,

RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kociba H.,

RA Knehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,

RA Schmitt L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Offield D., Fujino M., Aono H., Baldarelli R., Barsch G.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustigich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaeets P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,

RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,

RA Hayashizaki Y.,

RT "Functional annotation of a full-length mouse cDNA collection."

RL Nature 409:685-690(2001).

DR EMBL; AK01344; BAB27556.1;

DR MGD; MGI:1524219; Chst1.

SO SEQUENCE 114 AA; 13232 MW; ABA829F593B80892 CRC64;

Query Match 100.0%; Score 31; DB 11; Length 114;

Best Local Similarity 100.0%; Pred. No. 24;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6

DB 3 VREEDL 8

RESULT 2

Q9VMC2 PRELIMINARY; PRT; 183 AA.

AC Q9VMC2; 01-MAY-2000 (TREMBlrel. 13, Created)

DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)

DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)

DE Cg13765 protein.

GN Cg13765.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Phyllophaga; Drosophilidae; Drosophila.

NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;

RA MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.C.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Ffannkoch C., Baldwin D.,

RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasly E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,

RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durdin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,

RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glaeser K.,

RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasro P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,

RA Reiter K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spirdling A.C., Stapleton M., Strong R., Sun E.,

RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,

RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,

RT "The genome sequence of Drosophila melanogaster."

RL Science 287:2185-2195(2000).

DR EMBL: AE003614; AAF52399.1; -
 DR FlyBase: FBgn0031828; CG13765.
 SQ SEQUENCE 183 AA; 21351 MW; C353243BFAFEED8 CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 5; Length 183;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 75 VREEDL 80

RESULT 3
 0980N6 PRELIMINARY; PRT; 281 AA.

AC 0980N6
 DT 01-OCT-2001 (TREMBLrel. 18, Created)
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein SS00254.
 GN SS00254.

OS Sulfolobus solfataricus.
 OC Archaeae; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=2287;
 RN [1]

RC SEQUENCE FROM N.A.
 RP STRAIN-ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;
 RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
 De Moors A., Erasmo G., Fletcher C., Gordon P.M.K.,
 RA Helkamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2,"
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).

DR EMBL: AE006661; AAK40593.1; -
 DR InterPro: IPR002747; DUF62.
 DR InterPro: IPR001220; Lectin_legB.
 DR Pfam: PF01887; DUF62; 1.
 DR ProDom: PD013834; DUF62; 1.
 DR PROSITE: PS00307; LECTIN_LEGUME_BETA; UNKNOWN_1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 281 AA; 31911 MW; 23B2433BE3FE375F CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 17; Length 281;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 160 VREEDL 165

RESULT 4
 011316 PRELIMINARY; PRT; 331 AA.

AC 011316
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE B1-13 protein (Fragment).

OS Moluscum contagiosum virus subtype 1 (MCV1).
 OC Viruses; GSDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
 OC Moluscipoxvirus.
 OX NCBI_TaxID=10280;
 RN [1]

RP SEQUENCE FROM N.A.
 RA Moratilla M., Agromayor M., Nunez A., Funes J.M., Varas A.J.,

RA Lopez-Estebarez J.L., Esteban M., Martin-Gallardo A;
 RT "A Random DNA Sequencing, Computer-based Approach for the Generation
 RT of a Gene Map of Moluscum Contagiosum Virus,"
 RL Virus Genes 0:0-0(1997).

DR EMBL: U86899; AAB57946.1; -
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 331 AA; 37009 MW; E776B399590B97E CRC64;

Query Match
 Best Local Similarity 100.0%; Score 31; DB 12; Length 331;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 225 VREEDL 230

RESULT 5
 09Y5R3 PRELIMINARY; PRT; 386 AA.

AC 09Y5R3
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE N-acetylglucosamine 6-O-sulfotransferase (L-selectin ligand
 DE sulfotransferase GSF-3).
 GN GSF3.

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]

RC SEQUENCE FROM N.A.
 RP TISSUE-TONSIL;
 RX MEDLINE=99264336; PubMed=10330415;
 RA Bistrup A., Bhakta S., Lee J.K., Belov Y.Y., Gunn M.D., Zuo F.R.,
 RA Huang C.C., Kannagi R., Rosen S.D., Hemmerich S.;
 RT "Sulfotransferases of two specificities function in the reconstitution
 RT of high endothelial cell ligands for L-selectin,"
 RL J. Cell Biol. 145:899-910(1999).

[2]
 RC SEQUENCE FROM N.A.
 RP TISSUE-TONSIL;
 RC Hemmerich S., Lee J.-K., Bhakta S., Bistrup A., Ruddle N.R.,
 RA Rosen S.D.;
 RT "Chromosomal Localization and Genomic Organization for the
 RT Galactose/N-Acetylglucosamine/N-Acetylglucosamine 6-O-
 RT Sulfotransferase Gene Family,"
 RL Glycobiology 0:0-0(2001).

[3]
 RC SEQUENCE FROM N.A.
 RP MEDLINE=21332592; PubMed=11439191;
 RA Yeh J.-C., Hiraoaka N., Petryniak B., Nakayama J., Elies L.G.,
 RA Rabuka D., Hindsgrau O., Marth J.D., Lowe J.B., Fukuda M.;
 RT "Novel sulfated lymphocyte homing receptors and their control by a
 RT core1 extension beta1,3-N-acetylglucosaminyltransferase,"
 RL Cell 105:957-969(2001).

DR EMBL: AF131235; AAD33015.1; -
 DR EMBL: AF280088; AAG48246.1; -
 DR EMBL: AF149783; AAK48417.1; -
 DR InterPro: IPR001092; HLH_Basic.
 DR PROSITE: PS00038; HELIX_LOOP_HELIX; UNKNOWN_1.
 KW Lectin; Selection; Transferase.

Query Match
 Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
 |||||
 DB 274 VREEDL 279

```

RESULT 6
ID 090UP4 PRELIMINARY; PRT; 395 AA.
AC 090UP4;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE N-acetylglucosamine 6-O-sulfotransferase.
GN CHST5 OR I-GLCNAC-6-ST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=INTESTINE;
RA MEDLINE=99423499; PubMed=10491328;
RA Lee J.-K., Bhakta S., Rosen S.D., Hemmerlich S.;
RT "Cloning and Characterization of a Mammalian N-Acetylglucosamine-6-
RT Sulfotransferase that is Highly Restricted to Intestinal Tissue.";
RL Biochem. Biophys. Res. Commun. 263:543-549(1999).
DR EMBL; AF176841; AAD56003.1; -
DR EMBL; AF176840; AAD56002.1; -
DR MGD; MGI:1931825; Chst5.
KW Transferase.
SQ SEQUENCE 395 AA; 44537 MW; 3FDF71EA3ED383BE CRC64;

Query Match 100.0%; Score 31; DB 11; Length 395;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDEL 6
DB 271 VREDEL 276

RESULT 7
ID 043916 PRELIMINARY; PRT; 411 AA.
AC 043916;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-MAY-1999 (TREMBlrel. 10, Last annotation update)
DE Chondroitin 6-sulfotransferase.
GN CHST1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=97189336; PubMed=9117134;
RA Williams K.J.;
RT "Atherosclerosis: cell biology and lipoproteins.";
RL Curr. Opin. Lipidol. 7:0-0(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC Peng T., Tadas I., Williams K.J.;
RA Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RA MEDLINE=88070405; PubMed=9405439;
RA Fukuta M., Inazawa J., Torii T., Tsuzuki K., Shimada E., Habuchi O.;
RT "Molecular cloning and characterization of human keratan sulfate
RT Gal-6-sulfotransferase.";
RL J. Biol. Chem. 272:32321-32328(1997).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=DUBILICAL VEIN ENDOTHELIUM;
RA MEDLINE=99168906; PubMed=10049591;

```

```

RA Li X., Tedder T.F.;
RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
RT endothelial cells: cDNA cloning, expression, and chromosomal
RT localization.";
RL Genomics 55:345-347(1999).
DR EMBL; U65637; AAC28776.1; -
DR EMBL; AB003791; BAA24840.1; -
DR EMBL; AF090137; AAD19878.1; -
KW Transferase.
SQ SEQUENCE 411 AA; 46714 MW; F07D0A23B6338A09 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 411;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDEL 6
DB 300 VREDEL 305

RESULT 8
ID 09EOC0 PRELIMINARY; PRT; 411 AA.
AC 09EOC0;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Galactose 6-O-sulfotransferase GST-1.
GN CHST1 OR GST1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=MAMMARY GLAND.
RA Hemmerlich S., Lee J.-K., Bhakta S., Bistup A., Ruddle N.R.,
RA Rosen S.D.;
RT "Chromosomal localization and genomic organization of the galactose/N-
RT acetylglucosamine/N-acetylglucosamine 6-O-sulfotransferase gene
RT family.";
RL Glycobiology 11:75-87(2000).
DR EMBL; AF280087; AAG48245.1; -
DR MGD; MGI:1924219; Chst1.
KW Transferase.
SQ SEQUENCE 411 AA; 46903 MW; B1AE590EF5B9CBDC CRC64;

Query Match 100.0%; Score 31; DB 11; Length 411;
Best Local Similarity 100.0%; Pred. No. 98;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREDEL 6
DB 300 VREDEL 305

RESULT 9
ID 09RVO3 PRELIMINARY; PRT; 428 AA.
AC 09RVO3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein DR0969.
GN DR0969.
OS Deinococcus radiodurans.
OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RI;
RA MEDLINE=20036896; PubMed=10567266;

```


RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
 RA Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium *Deinococcus*
 RT *radiodurans* R1.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AF001949; AAF10548.1; -.
 DR TIGR: DR0969; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 428 AA; 46258 MW; 00ED2F1682C0DF45 CRC64;

Query Match 100.0%; Score 31; DB 16; Length 428;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 120 VRYEDL 125

RESULT 10

09UED5 PRELIMINARY; PRT; 484 AA.
 AC 09UED5;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE N-acetylglucosamine-6-O-sulfotransferase (GlcNAc6ST).
 GN GN6ST.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=9391845; PubMed=9722682;
 RA Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q.,
 RA Mitsuoka C., Kanagaki R., Habuchi O., Yokoyama I., Yamamura K.,
 RA Ozaki T., Nakagawara A., Kadomatsu K., Muramatsu T.;
 RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the
 RT biosynthesis of 6-sulfo sialyl Lewis X: Molecular cloning, chromosomal
 RT mapping, and expression in various organs and tumor cells.";
 RL J. Biochem. 124:670-678(1998).
 DR EMBL: AB014679; BAA34265.1; -.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 484 AA; 53641 MW; D077EBBD7645F2F0 CRC64;

Query Match 100.0%; Score 31; DB 4; Length 484;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 377 VRYEDL 382

RESULT 11

09JTN9 PRELIMINARY; PRT; 486 AA.
 AC 09JTN9;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE LD3505JP.
 GN CG9551.
 OS *Drosophila melanogaster* (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydriidae; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RA Stapleton M., Brickstein P., Hong L., Abdayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Fiske E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nuno J., Paclio J., Paragas V., Park S., Phoumenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Celisner S.;
 RL Submitted (Oct-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY058647; AAL13876.1; -.
 DR FLYBase: FBgn0031827; CG9551.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 486 AA; 53539 MW; BEDC96F096746EB CRC64;

Query Match 100.0%; Score 31; DB 5; Length 486;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 378 VRYEDL 383

RESULT 12

09Y6F2 PRELIMINARY; PRT; 530 AA.
 AC 09Y6F2; Q9GZNF;
 DT 01-NOV-1999 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Carbohydrate sulfotransferase 2 (N-acetylglucosamine
 DE 6-O-sulfotransferase).
 GN CHST2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE=99168906; PubMed=10049591;
 RA Li X., Tedder T.F.;
 RT "CHST1 and CHST2 sulfotransferases expressed by human vascular
 RT endothelial cells: cDNA cloning, expression, and chromosomal
 RT localization.";
 RL Genomics 55:345-347(1999).
 RN [2]
 RP SEQUENCE OF 48-530 FROM N.A.
 RC TISSUE-PLACENTA;
 RX MEDLINE=20498786; PubMed=11042394;
 RA Sakaguchi H., Kitagawa H., Sugahara K.;
 RT "Functional expression and genomic structure of human N-
 RT acetylglucosamine-6-O-sulfotransferase that transfers sulfate to b-N-
 RT acetylglucosamine at the nonreducing end of an N-acetylglucosamine
 RT sequence.";
 RL Biochim. Biophys. Acta 1523:269-276(2000).
 DR EMBL: AF083066; AAD20981.1; -.
 DR EMBL: AB021125; BAB16887.1; -.
 DR EMBL: AB021124; BAB16886.1; -.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KW Transferase.
 SQ SEQUENCE 530 AA; 57857 MW; A82CA227B9D5651B CRC64;

Query Match 100.0%; Score 31; DB 4; Length 530;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 423 VREEDL 428

RESULT 13

ID 088276 PRELIMINARY; PRT; 530 AA.

AC 088276;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE N-acetylglucosamine-6-O-sulfotransferase long form.
GN CHST2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=WHOLE EMBRYOS;
RX MEDLINE=98380482; PubMed=9712885;
RA Uchimura K., Muramatsu H., Kadomatsu K., Fan Q., Kurosawa N.,
RA Mitsuoka C., Kannagi R., Habuchi O., Muramatsu T.,
RT "Molecular cloning and characterization of an N-acetylglucosamine-6-O-sulfotransferase";
RL J. Biol. Chem. 273:22577-22583(1998).
DR EMBL; AB011452; BAA32138.1;
DR EMBL; AB011452; BAA32139.1;
DR MGI; MGI:1891160; Chst2.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.
KM TRANSFERENCE 530 AA; 57814 MW; A113E1B735C63EC CRC64;
SQ SEQUENCE

Query Match 100.0%; Score 31; DB 11; Length 530;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 423 VREEDL 428

RESULT 14

ID 09Y4C5 PRELIMINARY; PRT; 531 AA.

AC 09Y4C5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Long form of N-acetylglucosamine-6-O-sulfotransferase (G1CNA6ST).
GN G1CNA6ST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=98391845; PubMed=9722682;
RA Uchimura K., Muramatsu H., Kaname T., Ogawa H., Yamakawa T., Fan Q.,
RA Mitsuoka C., Kannagi R., Habuchi O., Yokoyama I., Yamamura K.,
RA Ozaki T., Nakagawara A., Kadomatsu K., Muramatsu T.,
RT "Human N-acetylglucosamine-6-O-sulfotransferase involved in the biosynthesis of 6-sulfo sialyl Lewis x: Molecular cloning, chromosomal mapping, and expression in various organs and tumor cells";
RL J. Biochem. 124:670-678(1998).
DR EMBL; AB014680; BAA34266.1;
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.

KW Transferase.
SQ SEQUENCE 531 AA; 58768 MW; D2EDB74E95B5162F CRC64;

Query Match 100.0%; Score 31; DB 4; Length 531;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 424 VREEDL 429

RESULT 15

ID 09Y158 PRELIMINARY; PRT; 597 AA.

AC 09Y158;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE BCDNA:GH02901.
GN BCDNA:GH02901 OR CG9009.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephyroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Rubin G.M., Wan K.H., Harvey D., Lewis S.E., Brokstein P., Tsang G.,
RA Agbayani A., Arcaina T.T., Baxter E., Blazej R.G., Butenoff C.,
RA Champe M., Chavez C., Chew M., Doyle C.M., Farfan D.E., Frise E.,
RA Galle R., George R.A., Harris N.L., Hoskins R.A., Evans-Holm M.,
RA Houston K.A., Hummasti S.R., Kim E., Li P., Mostrelet M., Pacle J.M.,
RA Park S., Seguela A., Sethi H., Snir E., Svirsks R.R., Weinburg T.,
RA Ceiniker S.E.;
RT "Full length Drosophila melanogaster cDNA sequence";
RL Submitted (APR-1999) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF145610; AAD38585.1;
DR HSSP: P08659; ILCI.
DR FlyBase: FBgn0027601; BCDNA:GH02901.
DR InterPro: IPR000873; AMP-bind.
DR Pfam: PF00501; AMP-binding; 1.
DR PROSITE: PS00455; AMP-BINDING; 1.
SQ SEQUENCE 597 AA; 65570 MW; 088010B38190DB22 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
|||||
DB 229 VREEDL 234

RESULT 16

ID 025212 PRELIMINARY; PRT; 631 AA.

AC 025212;
DT 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein HP0465.
GN HP0465.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=97394467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klenk H.-P., Gill S., Dougherty B.A.,

```

RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khalak H.G., Glodek A.,
RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gockayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Colton M.D., Weidman J.M., Fujii C., Bowman C., Watthey L., Wallin E.,
RA Hayes W.S., Borodovsky M., Kap P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
DR EMBL: AE000561; AAD07527.1; -.
DR TIGR: HP0465; -.
DR InterPro: IPR000834; Zn_carboxept.
DR PROSITE: PS00133; CARBOXYPEPT_2N_2; UNKNOWN_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 631 AA; 72171 MW; 81380827F060870B CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 16; Length 631;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 182 VREYDL 187

RESULT 17
Q9V884 PRELIMINARY; PRT; 649 AA.
ID Q9V884;
AC Q9V884;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)
DE CG18635 protein.
GN CG18635.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gockayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiri J.F., Adhyan A., An H.-T., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kelchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Liao P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puil V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,

```

```

RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
RA Svirskis R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR EMBL: AE003802; AAF57791.1; -.
DR Flybase: FBgn0034279; CG18635.
SQ SEQUENCE 649 AA; 74414 MW; 6471AC52B49C9AEB CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 5; Length 649;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 332 VREYDL 337

RESULT 18
Q98200 PRELIMINARY; PRT; 748 AA.
ID Q98200;
AC Q98200;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)
DE MC032L.
GN MC032L.
OS Moluscum contagiosum virus subtype 1 (MCV1).
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Molluscipoxvirus.
OX NCBI_TaxId=10280;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96325459; PubMed=8670425;
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RT "Genome sequence of a human tumorigenic poxvirus: Prediction of
RT specific host response-evasion genes."
RL Science 273:813-816(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,
RA Moss B.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U60315; AAC55160.1; -.
SQ SEQUENCE 748 AA; 84411 MW; 61A2741F62C0A58A CRC64;

Query Match
Best Local Similarity 100.0%; Score 31; DB 12; Length 748;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
DB 617 VREYDL 622

RESULT 19
Q45412 PRELIMINARY; PRT; 973 AA.
ID Q45412;
AC Q45412;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE F28B1.2 protein.
GN F28B1.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Peloderae; Caenorhabditis.
OX NCBI_TaxId=6239;

```

RN [1]
 RA SEQUENCE FROM N.A.
 RP Matthews L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99069613; PubMed=9851916;
 RA none;
 RT "Genome sequence of the nematode *C.elegans*: A platform for
 RT investigating biology.";
 RL Science 282:2012-2018(1998).
 DR EMBL: Z81517; CAB04209.1; -;
 SQ SEQUENCE 973 AA; 109292 MW; 74F2C8C63BA94067 CRC64;

Query Match 100.0%; Score 31; DB 5; Length 973;
 Best Local Similarity 100.0%; Pred. No. 2.5e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 828 VRYEDL 833

RESULT 20

OS4299 PRELIMINARY; PRT; 6260 AA.
 ID 054299
 AC 054299;
 DT 01-NOV-1996 (TREMblrel. 01, Created)
 DT 01-NOV-1996 (TREMblrel. 01, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Polypeptide synthase.
 GN RABC.
 OS Streptomyces hygroscopicus.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1912;
 RN [1]
 RA SEQUENCE FROM N.A.
 RP STRAIN=NRRL 5491;
 RC MEDLINE=9537374; PubMed=7644502;
 RX Schwecke T., Aparicio J.F., Molnar I., Koenig A., Khaw L.E.,
 RA Haydock S.F., Orligny M., Caffrey P., Cortes J., Lester J.B.,
 RA Boehm G.A., Staunton J., Leadlay P.F.,
 RT "The biosynthetic gene cluster for the polypeptide immunosuppressant
 RT rapamycin.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:7839-7843(1995).
 RN [2]
 RA SEQUENCE FROM N.A.
 RP STRAIN=NRRL 5491;
 RC MEDLINE=96186895; PubMed=8635730;
 RX Molnar I., Aparicio J.F., Haydock S.F., Khaw L.E., Schwecke T.,
 RA Koenig A., Staunton J., Leadlay P.F.;
 RT "Organisation of the biosynthetic gene cluster for rapamycin in
 RT Streptomyces hygroscopicus: analysis of genes flanking the polypeptide
 RT synthase.";
 RL Gene 169:1-7(1996).
 RN [3]
 RA SEQUENCE FROM N.A.
 RP STRAIN=NRRL 5491;
 RC MEDLINE=96186896; PubMed=8635756;
 RX Aparicio J.F., Molnar I., Schwecke T., Koenig A., Haydock S.F.,
 RA Khaw L.E., Staunton J., Leadlay P.F., Staunton J., Leadlay P.F.;
 RT "Organization of the biosynthetic gene cluster for rapamycin in
 RT Streptomyces hygroscopicus: analysis of the enzymatic domains in the
 RT modular polypeptide synthase.";
 RL Gene 169:9-16(1996).
 DR EMBL: XA60462.1; -;
 DR HSSP; P25715; IMIA.
 DR InterPro: IPR001227; Ac_transferase.
 DR InterPro: IPR002085; adh_zn_family.
 DR InterPro: IPR004410; Pabd.
 DR InterPro: IPR001899; Gram_pos_anchor.
 DR InterPro: IPR000794; Ketoacyl-synt.

DR InterPro: IPR003880; Ppantn_attach.
 DR InterPro: IPR002364; QOR_zeta_crystal.
 DR InterPro: IPR000634; S/T_dehydratase.
 DR Pfam: PF00698; Acyl_transf. 4.
 DR Pfam: PF00107; adh_zincf. 1.
 DR Pfam: PF00109; ketoacyl-synt. 4.
 DR Pfam: PF02801; ketoacyl-synt_C. 4.
 DR Pfam: PF00550; pp-binding. 4.
 DR TIGRfams: TIGR00128; fadD. 4.
 DR PROSITE: PS50075; ACP_DOMAIN. 4.
 DR PROSITE: PS00606; B_KETOACYL_SYNTHASE. 4.
 DR PROSITE: PS00165; DEHYDRATASE_SER_THR; UNKNOWN.1.
 DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN.1.
 DR PROSITE: PS00012; PHOSPHOPANTHETINE. 4.
 DR PROSITE: PS01162; QOR_ZETA_CRYSTAL. 1.
 DR Phosphopantetheine transferase.
 SQ SEQUENCE 6260 AA; 659509 MW; D76C002F91F2080 CRC64;

Query Match 100.0%; Score 31; DB 2; Length 6260;
 Best Local Similarity 100.0%; Pred. No. 1.9e+03;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 4365 VRYEDL 4370

RESULT 21

OS4299 PRELIMINARY; PRT; 194 AA.
 ID 0827B3
 AC 0827B3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Hypothetical protein PAE3336.
 GN PAE3336.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Thermoproteaceae; Pyrobaculum.
 OX NCBI_TaxID=13773;
 RN [1]
 RA SEQUENCE FROM N.A.
 RP STRAIN=IM2 / ATCC 51768 / DSM 7523;
 RX PubMed=11792865;
 RA Fitz-Gibbon S.T., Ladhner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*
 RT aerophilum.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 DR EMBL: AE009925; AAL64849.1; -;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 194 AA; 22556 MW; C3B5B464EEB58D0 CRC64;

Query Match 96.8%; Score 30; DB 17; Length 194;
 Best Local Similarity 83.3%; Pred. No. 72;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 DB 83 VRYEDL 88

RESULT 22

OS4299 PRELIMINARY; PRT; 199 AA.
 ID 09D806
 AC 09D806;
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE 2200002K05R1K protein.
 GN 2200002K05R1K.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia: Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 RN NCB1_TaxID=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shimagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana I.,
 Saito T., Okazaki Y., Gajobori T., Hono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashbunner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., Kang B., Kochia H.,
 Kuenl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Mashio T.,
 Sakai K., Okido T., Furuno M., Aoto H., Balarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hall D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kontuski S.,
 Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 DR EMBL: AK008621; BAB25783.1; -;
 DR MGD: MGI:1916387; 2200002K05Rik.
 DR InterPro: IPR003599; I9.
 DR SMART: SM00409; IG; 1.
 SQ SEQUENCE 199 AA; 22564 MW; 77400D2FF9AD4642 CRC64;

Query Match
 Best Local Similarity 96.8%; Score 30; DB 11; Length 199;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYEDL 6
 :|||||
 Db 140 IREYEDL 145

RESULT 23
 Q9VMC4 PRELIMINARY: PRT; 363 AA.
 AC Q9VMC4;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)
 DE CG9550 protein.
 GN CG9550.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCB1_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 Amaralides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 Brandon R.C., Rogers Y.H.C., Blazej R.G., Champe W., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 Bortone D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 Doadson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 Gloder A., Gong F., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 Nelson D.R., Nelson K.A., Nixon K.A., Nuskern D.R., Paclob J.M.,
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003614; AAF52397.1; -;
 DR FLYBase: FBgn0031826; CG9550.
 SQ SEQUENCE 363 AA; 41932 MW; C3CC09E608FF8AC CRC64;

Query Match
 Best Local Similarity 96.8%; Score 30; DB 5; Length 363;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYEDL 6
 :|||||
 Db 258 IREYEDL 263

RESULT 24
 O8TUC3 PRELIMINARY: PRT; 540 AA.
 AC O8TUC3;
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Methylamine methyltransferase corrinoid activation protein.
 GN RMA OR MA0150.
 OS Methanosarcina acetivorans.
 OC Archaea; Euryarchaeota; Methanococci; Methanosarcinales;
 OC Methanosarcinaceae; Methanosarcina.
 OX NCB1_TaxID=2214;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C2A / ATCC 35395 / DSM 2834;
 RX MEDLINE=21929760; PubMed=11932238;
 RA Galagan J.E., Nusbaum C., Roy A., Endrizzi M.G., Macdonald P.,
 Fitzhugh W., Calvo S., Engels R., Smirnov S., Atwood D., Brown A.,
 Allen N., Naylor J., Stange-Thomann N., DeArrelano K., Johnson R.,
 Linton L., McEwan P., McKernan K., Talmas J., Tirrell A., Ye W.,
 Zimmer A., Barber R.D., Cann I., Graham D.E., Girhane D.A., Guss A.M.,
 Hedderich R., Ingram-Smith C., Kuetner H.C., Krzycki J.A.,
 Leigh J.A., Li W., Liu J., Mukhopadhyay B., Reeve J.N., Smith K.,
 Springer T.A., Umayam L.A., White O., White R.H., de Macario E.C.,
 Ferry J.G., Jarrell K.F., Jing H., Macario A.J.L., Paulsen I.,
 Pritchett M., Sowers K.R., Swanson R.V., Zinder S.H., Lander E.,
 Metcalf W.W., Birren B.;
 RA "The genome of Methanosarcina acetivorans reveals extensive metabolic
 and physiological diversity.";
 RT Genome Res. 12:532-542(2002).
 RL EMBL: AE010672; AA003603.1; -;
 DR FLYBase: FBgn0031826; CG9550.
 SQ SEQUENCE 540 AA; 58864 MW; 7B3D3230355227B CRC64;

Query Match
 Best Local Similarity 96.8%; Score 30; DB 17; Length 540;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 326 IRYEDL 331

RESULT 25
Q9A7T6

PRELIMINARY; PRT; 579 AA.

AC Q9A7T6; 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Oxidoreductase, GMC family.
GN CC1634.

OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=155892;

RP SEQUENCE FROM N.A.
RN 11

RC STRAIN=ATCC 19089 / CB15;
MEDLINE=21173698; PubMed=11259647;
RA Nielsen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
Petrocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
Utterback T., Tran K., Wolf A., Vamathevan J., Emolaeva M., White O.,
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,
RT "Complete genome sequence of Caulobacter crescentus."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).

DR EMBL; AE005839; AAK23612.1; -
DR TIGR; CC1634; -
DR InterPro: IPR002005; NAD_binding.

KW Complete proteome.
SQ SEQUENCE 579 AA; 64397 MW; 7125C4DPAD618F10 CRC64;

Query Match 96.8%; Score 30; DB 16; Length 579;
Best Local Similarity 83.3%; Pred. No. 2.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 153 IRYEDL 158

RESULT 26
Q8SRF6

PRELIMINARY; PRT; 959 AA.

AC Q8SRF6; 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SEC31-like protein involved in vesicular transport from ER to
DE golgi.
GN ECU08.0210.

OS Eucephalitozoon cuniculi.
OC Eukaryota; Microsporidia; Unikaryoniidae; Eucephalitozoon.
OX NCBI_TaxID=6035;

RP SEQUENCE FROM N.A.
RN 11

RC STRAIN=GB-M1;
RA Genoscope; (APR-2001) to the EMBL/GenBank/DBJ databases.
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
RN 12

RP SEQUENCE FROM N.A.
RC STRAIN=GB-M1;
RX MEDLINE=21576510; PubMed=11719806;
RA Katinka M.D., Duprat S., Cornillot E., Metenier G., Thomarat F.,
Prenstier G., Barde V., Peyretailade E., Brottier P., Wincker P.,

RA Delbac F., El Alaoui H., Peyret P., Saurin W., Gouy M.,
RA Weissenbach J., Vivares C.P.;
RT "Genome sequence and gene compaction of the eukaryote parasite
RT Eucephalitozoon cuniculi."
RL Nature 414:450-453(2001).
DR EMBL; AL590448; CAD26327.1; -
SQ SEQUENCE 959 AA; 104621 MW; 6EC8782742424322 CRC64;

Query Match 96.8%; Score 30; DB 5; Length 959;
Best Local Similarity 83.3%; Pred. No. 4.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 584 IRYEDL 589

RESULT 27
O23559

PRELIMINARY; PRT; 1322 AA.

AC O23559; 01-JAN-1998 (TREMBlrel. 05, Created)
DT 01-JAN-1998 (TREMBlrel. 05, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Hypothetical 150.2 kDa protein.
GN AT4G17140.

OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.
RA Bevan M., Stiekema W., Murphy G., Wambutt R., Pohl T., Terry N.,
RA Kreis M., Kavanagh T., Entlan K.D., Rieger M., James R.,
RA Puidomenech P., Hatzopoulos P., Obermaier B., Duesterhoft A.,
RA Jones J., Palme K., Ansong W., Delserny M., Bancroft I., Meves H.W.,
RA Scheller C., Chaiwatzis N.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN 12

RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z97342; CAB10495.1; -
DR EMBL; AL161545; CAB80985.1; -
DR InterPro: IPR01849; PH.

DR Pfam; PF00169; PH.1.
DR SMART; SM00233; PH.1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
KW Hypothetical protein.
SQ SEQUENCE 1322 AA; 150221 MW; F70DE47DB21CEAE4 CRC64;

Query Match 96.8%; Score 30; DB 10; Length 1322;
Best Local Similarity 83.3%; Pred. No. 5.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 21 IRYEDL 26

RESULT 28
Q9A994

PRELIMINARY; PRT; 234 AA.

AC Q9A994; 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Hypothetical protein CC1100.
GN CC1100.

OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.

OX NCBI_TaxID=155892;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; PubMed=11259647;
RA Nleiman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA Deboy R.T., Dodson R.J., Durkin A.S., Gwin M.L., Haft D.H.,
RA Koloney J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,
RA Utterback T., Tran K., Wolf A., Yamathavan J., Ermolaeva M., White O.,
RA Salberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RI "Complete genome sequence of *Caulobacter crescentus*."
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005787; AAK23084.1; -
DR TIGR: CC1100; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 234 AA; 26770 MW; FF37A0FBA895638 CRC64;

Query Match 93.5%; Score 29; DB 16; Length 234;
Best Local Similarity 83.3%; Pred. No. 1.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYED 6
DB 146 VREYED 151

RESULT 29
O9JFS0 PRELIMINARY; PRT; 247 AA.
ID O9JFS0;
AC O9JFS0;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Secreted chemokine binding protein K2R (35 kDa chemokine binding
DE protein).
GN K2R OR CRBP.
OS Ectromelia virus.
OC Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC Orthopoxvirus.
OC NCBI_TaxID=12643;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MOSCOM;
RX MEDLINE=20192152; PubMed=10725549;
RA Chen N., Buller R.M.L., Wall E.M., Upton C.;
RT "Analysis of host response modifier ORFs of ectromelia virus, the
RT causative agent of mousepox."
RL Virus Res. 66:155-173(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HAMSTREAD;
RX MEDLINE=20411419; PubMed=10954546;
RA Smith V.P., Alcamí A.;
RT "Expression of secreted cytokine and chemokine inhibitors by
RT ectromelia virus."
RL J. Virol. 74:8460-8471(2000).
DR EMBL: AF012825; AAC99577.1; -
DR EMBL: AJ277111; CAC05575.1; -
DR InterPro: IPR003184; Orthopox_35KD.
DR Pfam: PF02250; Orthopox_35KD; 1.
SQ SEQUENCE 247 AA; 26733 MW; B84F26F55E82587C CRC64;

Query Match 93.5%; Score 29; DB 12; Length 247;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYED 6
DB 179 VREYED 184

RESULT 30
O93JF6 PRELIMINARY; PRT; 303 AA.
ID O93JF6;
AC O93JF6;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical protein STBAC16H6.16C.
GN STBAC16H6.16C OR SC05981 OR SCBAC16H6.16C.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Collins M.R., Harris D.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RX MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
RA Rabbinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)."
RL Nature 417:141-147(2002).
DR EMBL: AL596162; CAC44595.1; -
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam: PF00685; Sulfotransfer; 1.
KW Hypothetical protein.
SQ SEQUENCE 303 AA; 34181 MW; A0408E22B0DFF666 CRC64;

Query Match 93.5%; Score 29; DB 16; Length 303;
Best Local Similarity 83.3%; Pred. No. 2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREYED 6
DB 221 VREYED 226

RESULT 31
O926E4 PRELIMINARY; PRT; 355 AA.
ID O926E4;
AC O926E4;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Putative iron ABC transporter periplasmic solute-binding protein.
GN RH1430 OR SM20723.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OS Plasmid pSymB (megaplasmid 2).

CC Bacteria: Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Rhizobiaceae; Sinorhizobium.
 OX NCBI_TaxID=382;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=1021;
 RA MEDLINE=21365608; PubMed=11481431;
 RA Folan T.M., Weidner S., Wong K., Buhrmester J., Chain P.,
 RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowie A., Gouzy J.,
 RA Golding B., Pehler A.;
 RT "The complete sequence of the 1.683-kb pSYM megaplasmid from the N2-
 RT fixing endosymbiont Sinorhizobium meliloti."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).
 DR EMBL; AF603647; CAC49830.1;
 DR InterPro: IPR000567; SBP_Bacterial_1;
 DR Pfam: PF01547; SBP_Bacterial_1;
 KW Plasmid; Hypothetical protein; Complete proteome.
 SQ SEQUENCE 355 AA; 39502 MW; B5B3E6094D7ACB CRC64;

Query Match 93.5%; Score 29; DB 16; Length 355;
 Best Local Similarity 83.3%; Pred. No. 2.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYEDL 6
 DB 63 VYEDM 68

RESULT 32
 ID 081244 PRELIMINARY; PRT; 393 AA.
 AC 081244;
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Polygalacturonase precursor.
 GN Mpg1.
 OS Cucumis melo (Muskmelon).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids I; Cucurbitales; Cucurbitaceae; Cucumis.
 OX NCBI_TaxID=3656;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hadfield K.A., Rose J.K.C., Yaver D.S., Berka R.M., Bennett A.B.;
 RT "Polygalacturonase gene expression in ripe melon fruit supports a role
 RT for polygalacturonase in ripening-associated pectin disassembly."
 RL Plant Physiol. 116:0-0(1998)
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL; AF062465; AAC26510.1;
 DR InterPro: IPR000743; GH28.
 DR InterPro: IPR000408; Reg_chr_condens.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 DR PROSITE: PS00502; POLYGALACTURONASE; 1.
 DR PROSITE: PS00626; KCC1_2; UNKNOWN_1.
 KW Cell wall; Glycosidase; Hydrolase; Signal.
 FT SIGNAL 1
 FT CHAIN 1 393 POTENTIAL.
 SQ SEQUENCE 393 AA; 42145 MW; B94F5635D5E8AC7 CRC64;

Query Match 93.5%; Score 29; DB 10; Length 393;
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYEDL 6
 DB 329 VYEDI 334

RESULT 33
 ID 022818 PRELIMINARY; PRT; 405 AA.

AC 022818;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Putative polygalacturonase.
 GN AT2G43860.
 OS Arabidopsis thaliana (Mouse-ear cress).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 CC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RX MEDLINE=20083487; PubMed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
 RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gill J.E.,
 RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhaver G.P., Preuss D., Nietman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 RT thaliana."
 RL Nature 402:761-768(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES
 CC (POLYGALACTURONASES).
 DR EMBL; AC002333; AAB64020.1;
 DR InterPro: IPR000743; GH28.
 DR Pfam: PF00295; Glyco_hydro_28; 1.
 KW Cell wall; Glycosidase; Hydrolase
 SQ SEQUENCE 405 AA; 43462 MW; 556D6557044D195 CRC64;

Query Match 93.5%; Score 29; DB 10; Length 405;
 Best Local Similarity 83.3%; Pred. No. 2.7e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VYEDL 6
 DB 335 VYEDI 340

RESULT 34
 ID P96458 PRELIMINARY; PRT; 154 AA.
 AC P96458;
 DT 01-MAY-1997 (TrEMBLrel. 03, Created)
 DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE HSPR.
 GN HSPR.
 OS Streptomyces albus G.
 CC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1962;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97157943; PubMed=9004222;
 RA Grandvalet C., Servant P., Mazodier P.;
 RT "Distruption of hspR, the repressor gene of the dnaX operon in
 RT Streptomyces albus G.";
 RL Mol. Microbiol. 23:77-84(1997).
 CC -1- SIMILARITY: BELONGS TO THE MERR FAMILY OF TRANSCRIPTIONAL
 CC REGULATORS.
 DR EMBL; U43299; AAB48078.1;
 DR InterPro: IPR000551; HTH_Merr.
 DR Pfam: PF00376; merr; 1.
 DR SMART; SM00422; HTH_MERR; 1.

DR PROSITE: PS00552: HTH_MERR_FAMILY: 1.
 KM DNA-binding: Transcription regulation.
 SQ SEQUENCE 154 AA; 17400 MW; F5C64F168BCF30B8 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 2; Length 154;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
 :|||||
 Db 132 VREYDV 137

RESULT 35
 ID 028741 PRELIMINARY; PRT; 195 AA.
 AC 028741;
 DT 01-JUN-1998 (TREMBLrel. 05, Created)
 DT 01-JAN-1998 (TREMBLrel. 05, last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 GN Hypothetical protein AF1531.
 OS Archaeoglobus fulgidus.
 OC Archaea; Euryarchaeota; Archaeoglobi; Archaeoglobales;
 OC Archaeoglobaceae; Archaeoglobus.
 OX NCBI_TaxID=2234;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VC-16 / DSM 4304 / ATCC 49558;
 RX MEDLINE=98049343; PubMed=9389475;
 RA Kleink H.-P., Claydon R.A., Tomb J.-F., White O., Nelson K.E.,
 RA Ketchum K.A., Dodson R.J., Gwinn M., Hickey E.R., Peterson J.D.,
 RA Richardson D.D., Kellavag A.R., Graham D.E., Kyriades N.C.,
 RA Fleischmann R.D., Quackenbush J., Lee N.H., Sutton G.G., Gill S.,
 RA Kirsness E.F., Dougherty B.A., McInerney K., Adams M.D., Loftus B.,
 RA Peterson S., Reich C.I., McNeil L.K., Badger J.H., Glodex A., Zhou L.,
 RA Overbeek R., Gocayne J.D., Weidman U.F., McDonald L., Uterback T.,
 RA Cotton M.D., Spriggs T., Artlich P., Kalne B.P., Sykes S.W.,
 RA Sadow P.W., D'Andrea K.P., Bowman C., Fujii C., Garland S.A.,
 RA Mason T.M., Olsen G.J., Fraser C.M., Smith H.O., Woese C.R.,
 RA Venter J.C.;
 RT "The complete genome sequence of the hyperthermophilic, sulphate-
 RT reducing archaeon Archaeoglobus fulgidus.";
 RL Nature 390:364-370(1997).
 DR EMBL: AE000997; AAB8717.1; -.
 KM TIGR: AF1531; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 195 AA; 23194 MW; C244F95420565E2C CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 17; Length 195;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
 :|||||
 Db 87 LREYDL 92

RESULT 36
 ID 0976N0 PRELIMINARY; PRT; 228 AA.
 AC 0976N0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
 DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 GN Hypothetical protein ST0160.
 OS Stulobus tokodaii.
 OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 OC Sulfolobus.
 OX NCBI_TaxID=111955;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=JCM 10545 / ?;
 RX PubMed=11572479;
 RA Kawarabayashi Y., Hino Y., Horikawa H., Jin-no K., Takahashi M.,
 RA Sekine M., Baba S.-I., Anka A., Kosugi H., Hosoyama A., Fukui S.,
 RA Nagai Y., Nishijima K., Otsuka R., Nakazawa H., Takamiya M., Kato Y.,
 RA Yoshizawa T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Masuda S., Yanagii M., Nishimura M., Yamagishi A.,
 RA Oshima T., Kikuchi H.;
 RT "Complete genome sequence of an aerobic thermocidophilic
 RT Crenarchaeon, Sulfolobus tokodaii strain7.";
 RL DNA Res. 8:123-140(2001).
 DR EMBL: AP000981; BAB5116.1; -.
 DR InterPro: IPR000209; Peptidase_S8.
 DR PROSITE: PS00336; SUBTILASE_ASF; UNKNOWN_1.
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 228 AA; 27361 MW; 7BCE2DC618163CE6 CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 17; Length 228;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
 :|||||
 Db 212 VREYDL 217

RESULT 37
 ID 09CEP1 PRELIMINARY; PRT; 233 AA.
 AC 09CEP1;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 GN Regulatory protein.
 GN YSPD OR L11795.
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillales;
 OC Streptococcaceae; Lactococcus.
 OX NCBI_TaxID=1360;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=IL1403;
 RX MEDLINE=21235186; PubMed=11337471;
 RA Bojorin A., Wincker P., Manger S., Jallion O., Malarne K.,
 RA Weissenbach J., Ehrlich S.D., Sorokin A.;
 RT "The complete genome sequence of the lactic acid bacterium Lactococcus
 RT lactis ssp. lactis IL1403.";
 RL Genome Res. 11:731-753(2001).
 DR EMBL: AE006409; AAK05893.1; -.
 KW Complete proteome.
 SQ SEQUENCE 233 AA; 27077 MW; 483DAD4E8896A8BD CRC64;

Query Match
 Best Local Similarity 90.3%; Score 28; DB 16; Length 233;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
 :|||||
 Db 146 IREYDI 151

RESULT 38
 ID 09RC18 PRELIMINARY; PRT; 273 AA.
 AC 09RC18;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
 DE 01-MAR-2002 (TREMBLrel. 20, last annotation update)
 GN Chok.
 OS Bacillus sp. KFB-CO4.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;
 OC Bacillaceae; Bacillus.

OC NCBI_TaxID=111825;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KFB-CO4;
 RX MEDLINE=20422179; PubMed=10966383;
 RA Yoon H.G., Kim H.Y., Lim Y.H., Kim H.K., Shin D.H., Hong B.S.,
 RA Cho H.Y.;
 RT "Thermotable chitosanase from *Bacillus* sp. Strain CK4: cloning and
 RT expression of the gene and characterization of the enzyme.";
 RL Appl. Environ. Microbiol. 66:3727-3734(2000).
 DR EMBL: AF160195; AAF24188.1;
 DR InterPro: IPR000400; GH_46.
 DR Pfam: PF01374; Glyco_hydro_46; 1.
 DR PROSITE: P560000; CHITOSANASE_46.80; UNKNOWN_1.
 SQ SEQUENCE 273 AA; 30552 MW; C93916EC9DD27B9 CRC64;

Query Match 90.3%; Score 28; DB 2; Length 273;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 218 VRYDDL 223

RESULT 39
 O9YTO1 PRELIMINARY; PRT; 273 AA.

AC O9YTO1;
 DT 01-MAY-1999 (TREMblrel. 10, Created)
 DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
 DE Mitogen.
 OS Ateles herpesvirus 3.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 CC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=85618;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=73;
 RX MEDLINE=20091363; PubMed=10623770;
 RA Albrecht J.C.;
 RT "Primary structure of the Herpesvirus Ateles genome.";
 RL J. Virol. 74:1033-1037(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=73;
 RA Albrecht J.C., Fleckenstein B.;
 RT Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF083424; AAC95536.1;
 DR InterPro: IPR001213; MMTV_SAg.
 DR Pfam: PF01054; MMTV_SAg; 1.
 SQ SEQUENCE 273 AA; 30599 MW; EC0418C2DEA56F1A CRC64;

Query Match 90.3%; Score 28; DB 12; Length 273;
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 208 VRYDDL 213

RESULT 40
 O9ET84 PRELIMINARY; PRT; 278 AA.

AC O9ET84;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Chitosanase precursor.
 OS *Bacillus amyloliquefaciens*.
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillales;

OC Bacillaceae; Bacillus.
 OX NCBI_TaxID=1390;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Seki K., Kuriyama H., Uchida Y.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Seki K., Kuriyama H., Okuda T., Uchida Y.;
 RT "Molecular cloning of the gene encoding chitosanase from *Bacillus*
 RT *amyloliquefaciens* UTk.";
 RL (in) Unknown A. (eds.);
 RL Advances in Chitin Science 2, pp.284-289, Unknown Publisher (1997).
 DR EMBL: AB051574; BAB19276.1;
 DR HSSP: P33665; 1CHR.
 DR InterPro: IPR000400; GH_46.
 DR Pfam: PF01374; Glyco_hydro_46; 1.
 DR PROSITE: P560000; CHITOSANASE_46.80; UNKNOWN_1.
 KW SIGNAL.
 FT SIGNAL.
 SQ SEQUENCE 278 AA; 31404 MW; A742779DB91B51B CRC64;

Query Match 90.3%; Score 28; DB 2; Length 278;
 Best Local Similarity 83.3%; Pred. No. 3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 223 VRYDDL 228

RESULT 41
 O9I732 PRELIMINARY; PRT; 306 AA.

AC O9I732;
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 15, Last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)
 DE Hypothetical protein PA0100.
 GN PA0100.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 CC Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stever C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Blinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lartig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL: AE004448; AAG03490.1;
 KW Hypothetical protein, Complete proteome.
 SQ SEQUENCE 306 AA; 33723 MW; E9EF587A69005111 CRC64;

Query Match 90.3%; Score 28; DB 16; Length 306;
 Best Local Similarity 83.3%; Pred. No. 3.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
 Db 16 VRYDDL 21

RESULT 42
 O82330 PRELIMINARY; PRT; 333 AA.

```

AC 082330;
DT 01-NOV-1998 (TREMblrel. 08, Created)
DT 01-NOV-1998 (TREMblrel. 08, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Putative steroid sulfotransferase.
GN AT2G14920.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Bento M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldlyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Talon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Niernan W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005396; AAC61289.1; -
DR HSSP; P50224; 1CJM.
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ
SEQUENCE 333 AA; 38924 MW; B13B671253C619CF CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 10; Length 333;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDEL 6
DB 232 MREDEL 237

RESULT 43
ID 045982 PRELIMINARY; PRT; 346 AA.
AC 045982;
DT 01-JUN-1998 (TREMblrel. 06, Created)
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE Hypothetical 40.0 kDa protein.
GN ZC513.9.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RX Wu X., Le T.T.;
RT "The sequence of C. elegans cosmid ZC513.";
```

```

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; U53155; AAC48271.1; -
DR InterPro: IPR003002; 7TM_chemol.
DR InterPro: IPR000168; 7TM_nematode.
DR Pfam; PF01461; 7tm_4; 1.
KW Hypothetical protein.
SQ
SEQUENCE 346 AA; 40010 MW; 30A80441AA784868 CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 5; Length 346;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDEL 6
DB 79 MREDEL 84

RESULT 44
ID 09FL45 PRELIMINARY; PRT; 354 AA.
AC 09FL45;
DT 01-MAR-2001 (TREMblrel. 16, Created)
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Steroid sulfotransferase-like protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=98344145; PubMed=9679202;
RA Kaneko T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,
RA Tabata S.;
RT "Structural analysis of Arabidopsis thaliana chromosome 5. V. Sequence
RT features of the regions of 1,381,565 bp covered by twenty one
RT physically assigned P1 and TAC clones.";
RL DNA Res. 5:131-145(1998).
DR EMBL; AB010697; BAB11159.1; -
DR InterPro: IPR000863; Sulfotransferase.
DR Pfam; PF00685; Sulfotransferase.
DR ProDom; PD001218; Sulfotransferase; 1.
KW Transferase.
SQ
SEQUENCE 354 AA; 40887 MW; E540A1449966A36C CRC64;

Query Match
Best Local Similarity 90.3%; Score 28; DB 10; Length 354;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDEL 6
DB 249 MREDEL 254

RESULT 45
ID 0940P4 PRELIMINARY; PRT; 359 AA.
AC 0940P4;
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE AFG07010/M09.18.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
```

OC eutrosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,
 RA Bann J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kaniya A.,
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.,
 RT "Arabidopsis cDNA clones." to the EMBL/GenBank/DBJ databases.
 RL Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY054219; AL06879.1; -
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransferase.1.
 DR Prodom: PD001218; Sulfotransferase.1.
 SO SEQUENCE 359 AA; 41378 MW; 0614F7987C3F4A CRC64;
 Query Match 90.3%; Score 28; DB 10; Length 359;
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 254 LRYEDL 259
 RESULT 46
 O9R111 PRELIMINARY; PRT; 388 AA.
 AC O9R111
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
 DE L-selectin ligand sulfotransferase.
 GN CHST4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=99361934; PubMed=10435581;
 RA Hiraoka N., Petryniak B., Nakayama J., Tsuboi S., Suzuki M., Yeh J.C.,
 RA Iwata D., Tanaka T., Miyasaka M., Lowe J.B., Fukuda M.,
 RT "A novel, high endothelial venule-specific sulfotransferase expresses
 RT 6-sulfo sialyl Lewis(x), an L-selectin ligand displayed by CD34.".
 RL Immunity 11:79-89(1999).
 DR EMBL: AF109155; MAD45579.1; -
 DR MGD: MGI:1349479; Chst4.
 KW Lectin; Selectin; Transferase.
 SO SEQUENCE 388 AA; 44636 MW; 6D5371AFB6884ABE CRC64;
 Query Match 90.3%; Score 28; DB 11; Length 388;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 273 LRYEDL 278
 RESULT 47
 O9WUE5 PRELIMINARY; PRT; 388 AA.
 AC O9WUE5
 DT 01-NOV-1998 (TREMBlrel. 12, Created)
 DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE N-acetylglucosamine 6-O-sulfotransferase (Carbohydrate Chondroitin
 DE 6/eratan) sulfotransferase 4).

GN CHST4.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=TONSTL;
 RA MEDLINE=99264336; PubMed=10330415;
 RA Bistrup A., Bhakta S., Lee J.-K., Belov Y.Y., Gunn M.D., Zuo F.-R.,
 RA Huang C.-C., Kamaqi R., Rosen S.D., Hemmerich S.,
 RT "Sulfotransferases of two specificities function in the reconstitution
 RT of high endothelial cell ligands for L-selectin.".
 RL J. Cell Biol. 145:899-910(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=TONGUE;
 RA MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamazaki I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 RA Schraml L.M., Staubli F., Suzuki R., Tomita M., Wagner I., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombere P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlski S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection.".
 RL Nature 409:685-690(2001).
 DR EMBL: AF131236; MAD33016.1; -
 DR EMBL: AK009113; BAB26078.1; -
 DR MGD: MGI:1349479; Chst4.
 KW Transferase.
 SO SEQUENCE 388 AA; 44694 MW; D1E9D7796DF8574D CRC64;
 Query Match 90.3%; Score 28; DB 11; Length 388;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 OY 1 VRYEDL 6
 DB 273 LRYEDL 278
 RESULT 48
 O97B92 PRELIMINARY; PRT; 391 AA.
 AC O97B92
 DT 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Molybdenum cofactor biosynthesis protein moea.
 GN TY0565 OR TVG0553325.
 OS Thermoplasma volcanium.
 OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatales;
 OC Thermoplasmataceae; Thermoplasma.
 OX NCBI_TaxID=50339;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GSS1 / DSM 4299 / JCM 9571;
 RA MEDLINE=20570466; PubMed=11121031;
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,
 RA Kawashima-Onya Y., Matsunabe K., Yamazaki M., Kanehori K., Kawamoto T.,
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.,

RT "Archaeal adaptation to higher temperatures revealed by genomic
 RT sequence of *Thermoplasma volcanium*."
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).
 DR EMBL: AF000992; BAR59707.1; -.
 DR InterPro: IPR001453; MOCF_biosynth.
 DR InterPro: IPR005111; Moea_C.
 DR InterPro: IPR005110; Moea_N.
 DR Pfam: PF00994; MOCF_biosynth. 1.
 DR Pfam: PF03454; Moea_C; 1.
 DR Pfam: PF03453; Moea_N; 1.
 DR ProDom: PD002460; MOCF_biosynth. 1.
 DR Trigram: TRIGR00177; molyb_syn; 1.
 KM Complete proteome.
 SQ SEQUENCE 391 AA; 42773 MW; 79BC9BA181FE99B CRC64;

QY 1 VREEDL 6
 DB 56 VREEDV 61

Query Match 90.3%; Score 28; DB 17; Length 391;
 Best Local Similarity 83.3%; Pred. No. 4.3e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 49
 O9V129 PRELIMINARY; PRT; 458 AA.
 AC O9V129;
 DT 01-MAY-2000 (TREMblrel. 13, Created)
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
 DE Tlud related protein.
 GN PAB1955.
 OS Pyrococcus abyssi.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=29292;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ORSAY;
 RA Heilig R.;
 RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
 RT structure and evolution."
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ248284; CAB49522.1; -.
 DR InterPro: IPR002510; PmbA_Tlud.
 DR Pfam: PF01523; PmbA_Tlud; 1.
 KM Complete proteome.
 SQ SEQUENCE 458 AA; 51567 MW; 3B520B59A370DB4D CRC64;

QY 1 VREEDL 6
 DB 25 LREYEDL 30

Query Match 90.3%; Score 28; DB 17; Length 458;
 Best Local Similarity 83.3%; Pred. No. 5.2e+02;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 50
 O93N59 PRELIMINARY; PRT; 472 AA.
 AC O93N59;
 DT 01-DEC-2001 (TREMblrel. 19, Created)
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 55.0 kDa protein.
 OS *Coxiella burnetii*.
 OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
 OC *Coxiella* group; *Coxiella*.
 OX NCBI_TaxID=777;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Hoover T.A., Vodkin M.H., Williams J.C., Culp D.W., Thompson H.A.;
 RT "A chromosomal DNA deletion explains the phenotype of the *Coxiella*
 RT burnetii phase II variant."
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF387640; AAK71263.1; -.
 DR InterPro: IPR000863; Sulfotransferase.
 DR Pfam: PF00685; Sulfotransfer; 1.
 KM Hypothetical protein.
 SQ SEQUENCE 472 AA; 54997 MW; 6E1507D77A53FA03 CRC64;

QY 1 VREEDL 6
 DB 288 IREEDL 293

Query Match 90.3%; Score 28; DB 2; Length 472;
 Best Local Similarity 66.7%; Pred. No. 5.3e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Search completed: February 20, 2003, 13:33:48
 Job time : 32.5714 secs

This Page Blank (uspio)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:30:31 ; Search time 6.85714 Seconds

(without alignments)
25.745 Million cell updates/sec

Title: US-09-816-825-9

Perfect score: 31

Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapect 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 1000 summaries

Database : Issued Patents,AA:*

1: /cgn2_6/prodata/1/1aa/5A.COMB.pep:*
2: /cgn2_6/prodata/1/1aa/5B.COMB.pep:*
3: /cgn2_6/prodata/1/1aa/6A.COMB.pep:*
4: /cgn2_6/prodata/1/1aa/6B.COMB.pep:*
5: /cgn2_6/prodata/1/1aa/PCITUS.COMB.pep:*
6: /cgn2_6/prodata/1/1aa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	6	4	US-09-045-284A-9
2	31	100.0	386	4	US-09-045-284A-2
3	31	100.0	386	4	US-09-190-911-1
4	31	100.0	411	4	US-09-015-188-2
5	31	100.0	483	3	US-09-263-023-2
6	31	100.0	483	4	US-09-471-867-2
7	31	100.0	484	3	US-09-263-023-4
8	31	100.0	484	4	US-09-471-867-4
9	28	90.3	306	5	US-08-867-030B-13
10	28	90.3	306	2	US-08-867-030B-13
11	28	90.3	458	2	US-08-655-879-2
12	28	90.3	479	2	US-08-899-514-2
13	28	90.3	626	1	US-08-472-934-6
14	28	90.3	626	2	US-08-323-460A-6
15	28	90.3	626	2	US-08-461-146C-6
16	28	90.3	626	3	US-08-461-145C-6
17	28	90.3	626	4	US-09-423-890-6
18	28	90.3	626	4	US-09-423-890-12
19	28	90.3	626	4	US-08-628-829-10
20	27	87.1	295	2	US-08-411-607A-2
21	27	87.1	308	3	US-08-927-433-9
22	27	87.1	308	4	US-08-775-882-4
23	27	87.1	308	4	US-08-775-882-6
24	27	87.1	308	4	US-09-276-600-6
25	27	87.1	551	2	US-09-067-351-2
26	27	87.1	551	2	US-09-360-490-2
27	27	87.1	606	4	US-09-041-236-4

28	27	87.1	983	4	US-09-412-554A-2	Sequence 2, Appli
29	27	87.1	1119	4	US-09-396-651B-2	Sequence 2, Appli
30	26	83.9	416	1	US-08-117-083-60	Sequence 60, Appli
31	26	83.9	465	2	US-08-588-983-20	Sequence 20, Appli
32	26	83.9	465	2	US-08-588-983-22	Sequence 22, Appli
33	26	83.9	465	2	US-08-588-976-20	Sequence 20, Appli
34	26	83.9	465	2	US-08-588-976-22	Sequence 22, Appli
35	26	83.9	919	2	US-08-588-993-9	Sequence 9, Appli
36	26	83.9	919	2	US-08-588-993-12	Sequence 12, Appli
37	26	83.9	919	2	US-08-588-976-9	Sequence 9, Appli
38	26	83.9	919	2	US-08-588-976-12	Sequence 12, Appli
39	26	83.9	1174	1	US-08-040-751-3	Sequence 3, Appli
40	26	83.9	1174	1	US-08-291-368-2	Sequence 2, Appli
41	26	83.9	1174	1	US-08-962-190-2	Sequence 2, Appli
42	26	83.9	1174	5	PCR-US95-10310-2	Sequence 2, Appli
43	26	83.9	1174	6	5164180-4	Sequence 2, Appli
44	26	83.9	2548	4	US-09-172-422-1	Patent No. 5164180
45	25	80.6	103	4	US-09-370-253-14	Sequence 14, Appli
46	25	80.6	171	4	US-09-352-990-20	Sequence 20, Appli
47	25	80.6	182	2	US-08-698-805-4	Sequence 4, Appli
48	25	80.6	248	4	US-09-210-7248A-2	Sequence 4, Appli
49	25	80.6	249	1	US-07-687-072B-2	Sequence 2, Appli
50	25	80.6	249	1	US-08-466-444-2	Sequence 2, Appli
51	25	80.6	322	4	US-09-134-001C-3194	Sequence 3194, Ap
52	25	80.6	391	4	US-09-134-001C-4880	Sequence 4880, Ap
53	25	80.6	400	4	US-09-352-990-18	Sequence 18, Appli
54	25	80.6	410	4	US-09-352-990-16	Sequence 16, Appli
55	25	80.6	432	4	US-09-370-253-12	Sequence 12, Appli
56	25	80.6	473	4	US-09-637-118B-2	Sequence 2, Appli
57	25	80.6	476	3	US-09-188-579-114	Sequence 114, App
58	25	80.6	476	4	US-09-315-444-114	Sequence 114, App
59	25	80.6	476	4	US-09-134-218-6	Sequence 6, Appli
60	25	80.6	476	4	US-09-721-365-114	Sequence 114, App
61	25	80.6	487	1	US-08-485-107-2	Sequence 2, Appli
62	25	80.6	608	4	US-09-134-001C-3855	Sequence 3855, Ap
63	25	80.6	615	1	US-08-484-105-6	Sequence 6, Appli
64	25	80.6	615	1	US-08-484-106-6	Sequence 6, Appli
65	25	80.6	853	4	US-09-254-352B-17	Sequence 17, Appli
66	25	80.6	854	4	US-09-254-352B-18	Sequence 18, Appli
67	25	80.6	883	1	US-08-106-433A-2	Sequence 2, Appli
68	25	80.6	883	4	US-09-254-352B-16	Sequence 16, Appli
69	25	80.6	1529	4	US-09-215-664-1	Sequence 1, Appli
70	25	80.6	2233	2	US-08-568-853-1	Sequence 1, Appli
71	25	80.6	2233	2	US-08-568-853-2	Sequence 2, Appli
72	25	80.6	2233	3	US-08-987-439-1	Sequence 1, Appli
73	24	77.4	20	1	US-08-171-299B-11	Sequence 11, Appli
74	24	77.4	26	4	US-08-858-207A-466	Sequence 466, App
75	24	77.4	52	4	US-09-369-494-17	Sequence 17, Appli
76	24	77.4	52	4	US-09-358-569D-15	Sequence 15, Appli
77	24	77.4	93	4	US-08-938-165A-466	Sequence 466, App
78	24	77.4	94	4	US-08-035-006-21	Sequence 21, Appli
79	24	77.4	109	4	US-09-134-001C-5334	Sequence 5334, Ap
80	24	77.4	115	3	US-08-483-749A-6	Sequence 6, Appli
81	24	77.4	125	1	US-08-225-989-20	Sequence 20, Appli
82	24	77.4	125	1	US-08-570-923-20	Sequence 20, Appli
83	24	77.4	125	1	US-08-580-014-20	Sequence 20, Appli
84	24	77.4	125	4	US-09-079-785-20	Sequence 20, Appli
85	24	77.4	141	4	US-09-079-785-20	Sequence 20, Appli
86	24	77.4	141	4	US-09-188-930-327	Sequence 327, App
87	24	77.4	147	6	5248606-46	Patent No. 5248606
88	24	77.4	148	3	US-08-584-031-12	Sequence 12, Appli
89	24	77.4	160	4	US-09-134-001C-4327	Sequence 4327, Ap
90	24	77.4	198	1	US-08-642-255-111	Sequence 111, App
91	24	77.4	198	1	US-08-642-255-111	Sequence 111, App
92	24	77.4	198	1	US-08-397-633A-41	Sequence 41, Appli
93	24	77.4	198	1	US-08-397-633A-52	Sequence 52, Appli
94	24	77.4	203	4	US-09-134-001C-5696	Sequence 5696, Ap
95	24	77.4	215	1	US-08-225-989-23	Sequence 23, Appli
96	24	77.4	215	1	US-08-570-923-23	Sequence 23, Appli
97	24	77.4	215	4	US-09-079-785-23	Sequence 23, Appli
98	24	77.4	218	4	US-08-914-375C-6	Sequence 6, Appli
99	24	77.4	218	4	US-08-914-375C-11	Sequence 11, Appli
100	24	77.4	218	4	US-08-914-375C-12	Sequence 12, Appli
	24	77.4	218	4	US-08-914-375C-14	Sequence 14, Appli

685	22	71.0	593	6	5523211-2	Patent No. 5523211	758	22	71.0	868	1	US-08-251-937A-6	Sequence 6, Appl 1
686	22	71.0	595	6	5523211-3	Patent No. 5523211	759	22	71.0	869	1	US-08-212-133A-3	Sequence 3, Appl 1
687	22	71.0	596	4	US-09-000-145-2	Sequence 2, Appl 1	760	22	71.0	870	4	US-09-177-650-91	Sequence 91, Appl 1
688	22	71.0	608	4	US-09-095-385-4	Sequence 4, Appl 1	761	22	71.0	872	4	US-09-177-650-7	Sequence 7, Appl 1
689	22	71.0	609	1	US-07-798-776-2	Sequence 2, Appl 1	762	22	71.0	874	3	US-08-823-110-1	Sequence 1, Appl 1
690	22	71.0	609	3	US-08-251-288A-2	Sequence 2, Appl 1	763	22	71.0	908	3	US-08-604-298-1	Sequence 1, Appl 1
691	22	71.0	609	3	US-09-298-819A-2	Sequence 2, Appl 1	764	22	71.0	908	3	US-08-699-1103B-12	Sequence 12, Appl 1
692	22	71.0	609	4	US-09-586-563C-2	Sequence 2, Appl 1	765	22	71.0	908	4	US-08-855-146-2	Sequence 2, Appl 1
693	22	71.0	609	4	US-09-586-562C-2	Sequence 2, Appl 1	766	22	71.0	908	4	US-09-229-051-12	Sequence 12, Appl 1
694	22	71.0	613	3	US-08-622-740-6	Sequence 6, Appl 1	767	22	71.0	912	4	US-08-617-785-2	Sequence 2, Appl 1
695	22	71.0	613	3	US-08-440-689-6	Sequence 6, Appl 1	768	22	71.0	912	4	US-09-641-315-2	Sequence 2, Appl 1
696	22	71.0	613	4	US-09-122-399-6	Sequence 6, Appl 1	769	22	71.0	912	5	PCT-US81-09442-19	Sequence 19, Appl 1
697	22	71.0	613	4	US-08-447-985-11	Sequence 11, Appl 1	770	22	71.0	918	2	US-08-843-530B-35	Sequence 35, Appl 1
698	22	71.0	614	3	US-08-622-740-8	Sequence 8, Appl 1	771	22	71.0	933	3	US-08-313-200-1	Sequence 1, Appl 1
699	22	71.0	614	3	US-08-440-689-8	Sequence 8, Appl 1	772	22	71.0	933	5	PCT-US93-03837-1	Sequence 1, Appl 1
700	22	71.0	614	4	US-09-122-399-8	Sequence 8, Appl 1	773	22	71.0	934	1	US-08-446-486-7	Sequence 7, Appl 1
701	22	71.0	614	4	US-08-447-985-13	Sequence 13, Appl 1	774	22	71.0	934	1	US-08-463-308-7	Sequence 7, Appl 1
702	22	71.0	615	3	US-09-042-426-10	Sequence 10, Appl 1	775	22	71.0	969	1	US-07-671-817A-4	Sequence 4, Appl 1
703	22	71.0	615	3	US-09-291-238-10	Sequence 10, Appl 1	776	22	71.0	969	1	US-07-671-817A-6	Sequence 6, Appl 1
704	22	71.0	615	4	US-09-330-760-10	Sequence 10, Appl 1	777	22	71.0	979	4	US-08-878-474-5	Sequence 5, Appl 1
705	22	71.0	615	4	US-09-338-473-10	Sequence 10, Appl 1	778	22	71.0	990	4	US-09-627-376-7	Sequence 7, Appl 1
706	22	71.0	615	4	US-09-330-737-10	Sequence 10, Appl 1	779	22	71.0	993	4	US-08-626-687B-30	Sequence 30, Appl 1
707	22	71.0	615	4	US-09-329-169-10	Sequence 10, Appl 1	780	22	71.0	1017	4	US-09-134-001C-3542	Sequence 3542, Ap
708	22	71.0	615	4	US-09-320-714A-10	Sequence 10, Appl 1	781	22	71.0	1040	4	US-08-961-083C-118	Sequence 118, Ap
709	22	71.0	615	4	US-09-328-826-10	Sequence 10, Appl 1	782	22	71.0	1042	4	US-09-387-695-2	Sequence 2, Appl 1
710	22	71.0	615	4	US-08-729-601A-2	Sequence 2, Appl 1	783	22	71.0	1065	3	US-08-630-172-9	Sequence 9, Appl 1
711	22	71.0	617	5	PCT-US96-03916-64	Sequence 64, Appl 1	784	22	71.0	1065	3	US-08-630-172-9	Sequence 9, Appl 1
712	22	71.0	625	4	US-08-959-004-10	Sequence 10, Appl 1	785	22	71.0	1090	6	US-09-375-419-9	Sequence 6, Appl 1
713	22	71.0	648	3	US-08-810-720-13	Sequence 13, Appl 1	786	22	71.0	1129	6	PCT-US93-03275-6	Patent No. 5164180
714	22	71.0	649	4	US-08-800-291B-5	Sequence 5, Appl 1	787	22	71.0	1155	1	US-08-349-867-19	Sequence 19, Appl 1
715	22	71.0	649	4	US-08-800-291B-6	Sequence 6, Appl 1	788	22	71.0	1155	1	US-08-349-867-33	Sequence 33, Appl 1
716	22	71.0	649	4	US-09-066-047-5	Sequence 5, Appl 1	789	22	71.0	1155	1	US-08-339-476-19	Sequence 19, Appl 1
717	22	71.0	650	4	US-08-800-291B-4	Sequence 4, Appl 1	790	22	71.0	1155	1	US-08-446-486-2	Sequence 2, Appl 1
718	22	71.0	657	4	US-09-370-368-7	Sequence 7, Appl 1	791	22	71.0	1155	1	US-07-951-715A-9	Sequence 9, Appl 1
719	22	71.0	669	3	US-08-704-711A-3	Sequence 3, Appl 1	792	22	71.0	1155	1	US-08-463-308-2	Sequence 2, Appl 1
720	22	71.0	669	4	US-09-521-220-3	Sequence 3, Appl 1	793	22	71.0	1155	2	US-08-598-305A-19	Sequence 19, Appl 1
721	22	71.0	669	4	US-09-391-104-29	Sequence 29, Appl 1	794	22	71.0	1155	2	US-08-598-305A-33	Sequence 33, Appl 1
722	22	71.0	677	4	US-08-522-269B-3	Sequence 3, Appl 1	795	22	71.0	1155	2	US-08-639-923A-19	Sequence 19, Appl 1
723	22	71.0	677	4	US-09-294-923-3	Sequence 3, Appl 1	796	22	71.0	1155	2	US-08-459-448A-9	Sequence 9, Appl 1
724	22	71.0	688	1	US-09-141-017-8	Sequence 8, Appl 1	797	22	71.0	1155	3	US-08-459-448A-9	Sequence 9, Appl 1
725	22	71.0	695	1	US-07-671-817A-5	Sequence 5, Appl 1	798	22	71.0	1155	3	US-09-021-203-2	Sequence 2, Appl 1
726	22	71.0	720	4	US-09-480-921B-8	Sequence 8, Appl 1	799	22	71.0	1155	3	US-08-459-444-9	Sequence 9, Appl 1
727	22	71.0	724	3	US-09-121-964-1	Sequence 1, Appl 1	800	22	71.0	1155	3	US-08-459-444-9	Sequence 9, Appl 1
728	22	71.0	742	4	US-09-215-664-12	Sequence 12, Appl 1	801	22	71.0	1155	3	US-09-053-549-4	Sequence 4, Appl 1
729	22	71.0	743	2	US-09-012-030-2	Sequence 2, Appl 1	802	22	71.0	1155	3	US-09-053-549-6	Sequence 6, Appl 1
730	22	71.0	743	2	US-08-530-454-2	Sequence 2, Appl 1	803	22	71.0	1155	4	US-09-547-422-9	Sequence 9, Appl 1
731	22	71.0	746	3	US-08-434-000A-4	Sequence 4, Appl 1	804	22	71.0	1155	5	PCT-US85-05431-19	Sequence 19, Appl 1
732	22	71.0	746	4	US-09-312-157-4	Sequence 4, Appl 1	805	22	71.0	1155	6	5254799-4	Patent No. 5254799
733	22	71.0	771	3	US-09-121-964-9	Sequence 9, Appl 1	806	22	71.0	1156	4	US-09-178-252-15	Sequence 15, Appl 1
734	22	71.0	815	1	US-08-123-520C-9	Sequence 9, Appl 1	807	22	71.0	1156	1	US-07-828-788A-14	Sequence 14, Appl 1
735	22	71.0	817	1	US-08-381-931B-2	Sequence 2, Appl 1	808	22	71.0	1155	1	US-08-239-476-33	Sequence 33, Appl 1
736	22	71.0	820	1	US-08-291-896-2	Sequence 2, Appl 1	809	22	71.0	1155	1	US-08-356-034-4	Sequence 4, Appl 1
737	22	71.0	820	2	US-08-485-278-2	Sequence 2, Appl 1	810	22	71.0	1155	2	US-08-639-923A-33	Sequence 33, Appl 1
738	22	71.0	827	4	US-08-659-286-11	Sequence 11, Appl 1	811	22	71.0	1155	3	US-08-933-891-4	Sequence 4, Appl 1
739	22	71.0	827	4	US-09-469-253-11	Sequence 11, Appl 1	812	22	71.0	1155	4	US-09-176-320-6	Sequence 6, Appl 1
740	22	71.0	827	4	US-09-642-146-11	Sequence 11, Appl 1	813	22	71.0	1155	5	PCT-US92-11337-14	Sequence 14, Appl 1
741	22	71.0	838	4	US-09-315-794-62	Sequence 52, Appl 1	814	22	71.0	1155	5	PCT-US95-05431-33	Sequence 33, Appl 1
742	22	71.0	838	4	US-09-369-341-62	Sequence 52, Appl 1	815	22	71.0	1155	6	5188960-4	Patent No. 5188960
743	22	71.0	838	4	US-09-564-805-229	Sequence 229, App	816	22	71.0	1156	2	US-08-789-078-2	Sequence 2, Appl 1
744	22	71.0	844	1	US-07-646-337B-2	Sequence 2, Appl 1	817	22	71.0	1170	2	US-08-752-633-2	Sequence 2, Appl 1
745	22	71.0	853	3	US-08-451-946B-6	Sequence 6, Appl 1	818	22	71.0	1170	2	US-08-476-062A-42	Sequence 42, Appl 1
746	22	71.0	853	3	US-08-446-938B-6	Sequence 6, Appl 1	819	22	71.0	1170	5	PCT-US95-04886-2	Sequence 2, Appl 1
747	22	71.0	853	3	US-08-311-703A-6	Sequence 6, Appl 1	820	22	71.0	1170	5	PCT-US96-01314-42	Sequence 42, Appl 1
748	22	71.0	853	3	US-08-446-939B-6	Sequence 6, Appl 1	821	22	71.0	1171	1	US-07-828-788A-6	Sequence 6, Appl 1
749	22	71.0	853	3	US-09-183-543-6	Sequence 6, Appl 1	822	22	71.0	1171	1	US-08-278-685-2	Sequence 2, Appl 1
750	22	71.0	853	3	US-08-699-103B-10	Sequence 10, Appl 1	823	22	71.0	1171	1	US-08-277-721-2	Sequence 2, Appl 1
751	22	71.0	853	4	US-08-446-936A-6	Sequence 6, Appl 1	824	22	71.0	1171	1	US-08-277-721-4	Sequence 4, Appl 1
752	22	71.0	853	4	US-09-229-059-10	Sequence 10, Appl 1	825	22	71.0	1171	1	US-08-602-737-4	Sequence 4, Appl 1
753	22	71.0	853	5	PCT-US92-09326-2	Sequence 2, Appl 1	826	22	71.0	1171	4	US-09-001-882-4	Sequence 4, Appl 1
754	22	71.0	854	4	US-09-105-058C-27	Sequence 27, Appl 1	827	22	71.0	1171	5	PCT-US92-11337-6	Sequence 6, Appl 1
755	22	71.0	863	4	US-09-238-303-11	Sequence 11, Appl 1	828	22	71.0	1176	1	US-07-828-788A-12	Sequence 12, Appl 1
756	22	71.0	866	4	US-09-134-001C-4930	Sequence 4930, Ap	829	22	71.0	1176	1	US-08-446-486-6	Sequence 6, Appl 1
757	22	71.0	868	1	US-07-864-004B-6	Sequence 6, Appl 1	830	22	71.0	1176	1	US-08-157-563A-10	Sequence 10, Appl 1


```
977 21 67.7 32 2 US-08-140-137A-46 Sequence 46, Appl
978 21 67.7 43 1 US-07-998-003A-73 Sequence 73, Appl
979 21 67.7 43 1 US-08-453-274B-73 Sequence 73, Appl
980 21 67.7 43 1 US-08-453-695A-73 Sequence 73, Appl
981 21 67.7 43 1 US-08-268-161A-73 Sequence 73, Appl
982 21 67.7 43 2 US-08-453-702A-73 Sequence 73, Appl
983 21 67.7 43 4 US-09-099-639-73 Sequence 73, Appl
984 21 67.7 43 5 PCT-US93-12588-73 Sequence 73, Appl
985 21 67.7 43 5 PCT-US95-08071-73 Sequence 73, Appl
986 21 67.7 52 2 US-08-209-521-16 Sequence 16, Appl
987 21 67.7 52 4 US-08-961-810-126 Sequence 126, App
988 21 67.7 52 4 US-08-352-902D-126 Sequence 126, App
989 21 67.7 53 4 US-08-851-843A-20 Sequence 20, Appl
990 21 67.7 53 4 US-08-854-050-20 Sequence 20, Appl
991 21 67.7 53 4 US-09-430-323-20 Sequence 20, Appl
992 21 67.7 59 2 US-08-633-879C-18 Sequence 18, Appl
993 21 67.7 61 6 5320958-19 Patent No. 5320958
994 21 67.7 64 2 US-08-209-521-14 Sequence 14, Appl
995 21 67.7 64 2 US-08-209-521-15 Sequence 15, Appl
996 21 67.7 64 4 US-08-961-810-124 Sequence 124, App
997 21 67.7 64 4 US-08-961-810-125 Sequence 125, App
998 21 67.7 64 4 US-08-352-902D-124 Sequence 124, App
999 21 67.7 64 4 US-08-352-902D-125 Sequence 125, App
1000 21 67.7 69 4 US-09-134-001C-4746 Sequence 4746, Ap
```

ALIGNMENTS

```
RESULT 1
US-09-045-284A-9
; Sequence 9, Application US/09045284A
; Patent No. 6265192
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107U01
; CURRENT APPLICATION NUMBER: US/09/045,284A
; CURRENT FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-9
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 6;
Pred. No. 1.9e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
LENGTH: 386
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-045-284A-2
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 274 VRYEDL 279
```

```
RESULT 3
US-09-190-911-1
; Sequence 1, Application US/09190911
; Patent No. 6365365
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Tangemann, Kirsten
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
; FILE REFERENCE: 6510-107CIP
; CURRENT APPLICATION NUMBER: US/09/190,911
; CURRENT FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/045,284
; EARLIER FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 386
; TYPE: PRT
; ORGANISM: H. sapiens
US-09-190-911-1
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 386;
Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 274 VRYEDL 279
```

```
RESULT 4
US-09-015-188-2
; Sequence 2, Application US/09015188C
; Patent No. 6399358
; GENERAL INFORMATION:
; APPLICANT: Williams, Kevin J
; APPLICANT: Tabas, Ira
; TITLE OF INVENTION: A Human Gene Encoding Human Chondroitin
; FILE REFERENCE: JEPF-0231
; CURRENT APPLICATION NUMBER: US/09/015,188C
; CURRENT FILING DATE: 1998-01-29
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 411
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-015-188-2
Query Match
Best Local Similarity 100.0%; Score 31; DB 4; Length 411;
Pred. No. 30;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
Db 300 VRYEDL 305
```

RESULT 5

US-09-263-023-2
Sequence 2, Application US/09263023
Patent No. 6037159

GENERAL INFORMATION:

APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji

APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND

FILE REFERENCE: TOYAMA1.001AUS

CURRENT APPLICATION NUMBER: US/09/263,023

EARLIER FILING DATE: 1999-03-05

EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 483

TYPE: PRT

ORGANISM: Mus musculus

US-09-263-023-2

Query Match 100.0%; Score 31; DB 3; Length 483;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

Db 376 VRYEDL 381

RESULT 6

US-09-471-867-2
Sequence 2, Application US/09471867
Patent No. 6455289

GENERAL INFORMATION:

APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji

APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND

FILE REFERENCE: TOYAMA1.001AUS

CURRENT APPLICATION NUMBER: US/09/471,867

EARLIER FILING DATE: 1999-12-23

EARLIER APPLICATION NUMBER: US 09/263,023

EARLIER FILING DATE: 1999-03-05

EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 2

LENGTH: 483

TYPE: PRT

ORGANISM: Mus musculus

US-09-471-867-2

Query Match 100.0%; Score 31; DB 4; Length 483;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
Db 376 VRYEDL 381

RESULT 7
US-09-263-023-4
Sequence 4, Application US/09263023
Patent No. 6037159

GENERAL INFORMATION:

APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji

APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND

FILE REFERENCE: TOYAMA1.001AUS

CURRENT APPLICATION NUMBER: US/09/263,023

EARLIER FILING DATE: 1999-03-05

EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 484

TYPE: PRT

ORGANISM: Homo sapiens

US-09-263-023-4

Query Match 100.0%; Score 31; DB 3; Length 484;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6

Db 377 VRYEDL 382

RESULT 8

US-09-471-867-4
Sequence 4, Application US/09471867
Patent No. 6455289

GENERAL INFORMATION:

APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji

APPLICANT: Habuchi, Osami
APPLICANT: Muramatsu, Takashi

TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-SULFOTRANSFERASE AND

FILE REFERENCE: TOYAMA1.001AUS

CURRENT APPLICATION NUMBER: US/09/471,867

EARLIER FILING DATE: 1999-12-23

EARLIER APPLICATION NUMBER: US 09/263,023

EARLIER FILING DATE: 1999-03-05

EARLIER APPLICATION NUMBER: JP 10-54007

EARLIER FILING DATE: 1998-06-24

NUMBER OF SEQ ID NOS: 10

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 4

LENGTH: 484

TYPE: PRT

ORGANISM: Homo sapiens

US-09-471-867-4

Query Match 100.0%; Score 31; DB 4; Length 484;

Best Local Similarity 100.0%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 377 VRYEDL 382

RESULT 9

US-08-867-030B-13
; Sequence 13, Application US/08867030B
; Patent No. 5948900
; GENERAL INFORMATION:
; APPLICANT: Yotner et al.
; TITLE OF INVENTION: Streptococcus pneumoniae
; TITLE OF INVENTION: Capsular Polysaccharide Genes and Flanking Regions
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Benjamin Aaron Adler, Ph.D., J.D.
; STREET: 8011 Candle Lane
; CITY: Houston
; STATE: TX
; COUNTRY: USA
; ZIP: 77071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch floppy
; COMPUTER: Apple Macintosh
; OPERATING SYSTEM: Macintosh
; SOFTWARE: Microsoft Word for Macintosh
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/867,030B
; FILING DATE: June 2, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/243,546
; FILING DATE: May 16, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Benjamin Aaron Adler, Ph.D., J.D.
; REGISTRATION NUMBER: 35,423
; REFERENCE/DOCKET NUMBER: D5923
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 777-2321
; TELEFAX: (713) 777-6908
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE:
; DESCRIPTION: protein
; HYPOTHETICAL: no
; ANTI-SENSE: no
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; IMMEDIATE SOURCE:
; POSITION IN GENE:
US-08-867-030B-13

Query Match

Best Local Similarity 90.3%; Score 28; DB 2; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 165 VRYEDV 170

RESULT 10
PCT-US95-06119-13
; Sequence 13, Application PC/TUS9506119
; GENERAL INFORMATION:
; APPLICANT:

TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE CAPSULAR
; TITLE OF INVENTION: POLYSACCHARIDE GENES AND FLANKING REGIONS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: United States of America
; ZIP: 77210

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
SOFTWARE: Patent Release #1.0, Version
SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/06119
; FILING DATE: CONCURRENTLY HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/243,546
; FILING DATE: 16-MAY-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Parker, David L.
; REGISTRATION NUMBER: 32,165
; REFERENCE/DOCKET NUMBER: AMCY018P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 306 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-06119-13

Query Match

Best Local Similarity 90.3%; Score 28; DB 5; Length 306;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 165 VRYEDV 170

RESULT 11

US-08-655-878-2
; Sequence 2, Application US/08655878
; Patent No. 5827713
; GENERAL INFORMATION:
; APPLICANT: FUKUTA, MASAKAZU
; TITLE OF INVENTION: DNA CODING FOR SULFOTRANSFERASE
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE:
; STREET:
; CITY:
; STATE:
; COUNTRY:
; ZIP:

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/655,878

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME:
REGISTRATION NUMBER:
REFERENCE/DOCKET NUMBER:
TELECOMMUNICATION INFORMATION:
TELEPHONE:
TELEFAX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 458
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-655-878-2

Query Match 90.3%; Score 28; DB 2; Length 458;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 348 VRYEDV 353

RESULT 12
US-08-899-514-2
Sequence 2, Application US/08899514
Patent No. 5910561
GENERAL INFORMATION:
APPLICANT: HABUCHI, OSAMI
APPLICANT: FUKUTA, MASAKAZU
TITLE OF INVENTION: POLYPEPTIDE OF GLYCOSAMINOGLYCAN
TITLE OF INVENTION: SULFONAMIDASE ORIGINATING FROM HUMAN AND DNA CODING
TITLE OF INVENTION: FOR THE POLYPEPTIDE
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: KNOBE, MARTENS, OLSON & BEAR, LLP
STREET: 620 NEWPORT CENTER DRIVE, SIXTEENTH FLOOR
CITY: NEWPORT BEACH
STATE: CALIFORNIA
COUNTRY: US
ZIP: 92660
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,514
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: DANIEL E. ALTMAN
REGISTRATION NUMBER: 34,115
REFERENCE/DOCKET NUMBER: TOYAM21.001AUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 714 760 0404
TELEFAX: 714 760 9502
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 479
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-514-2

Query Match 90.3%; Score 28; DB 2; Length 479;
Best Local Similarity 83.3%; Pred. No. 1.7e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 369 VRYEDV 374

RESULT 13
US-08-472-934-6
Sequence 6, Application US/08472934
Patent No. 5753446
GENERAL INFORMATION:
APPLICANT: JOHNSON, GARY L.
TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL
NUMBER OF SEQUENCES: 12
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lathive and Cockfield
STREET: 60 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/472,934
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/440,421
FILING DATE: 15-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/354,516
FILING DATE: 21-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 5,405,941
FILING DATE: 15-APR-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
ATTORNEY/AGENT INFORMATION:
NAME: Giulio A. DeConti, Jr. Esq.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: CPT-004DVC2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-472-934-6

Query Match 90.3%; Score 28; DB 1; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6

Db 64 VREDDV 69

RESULT 14
US-08-323-460A-6
Sequence 6, Application US/08323460A

Patent No. 5854043

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHOD AND PRODUCT FOR REGULATING CELL

RESPONSE TO EXTERNAL SIGNALS

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

SHERIDAN ROSS & MCINTOSH

STREET: 1700 LINCOLN STREET, SUITE 3500

CITY: DENVER

STATE: CO

COUNTRY: USA

ZIP: 80203

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323,460A

FILING DATE: 14-OCT-1994

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 14-APR-1993

ATTORNEY/AGENT INFORMATION:

NAME: KOVARIK, JOSEPH E.

REGISTRATION NUMBER: 33,005

REFERENCE/DOCKET NUMBER: 2879-1-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: 303/863-9700

TELEFAX: 303/863-0223

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-323-460A-6

Query Match 90.3%; Score 28; DB 2; Length 626;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDDV 6

Db 64 VREDDV 69

RESULT 15
US-08-461-146C-6
Sequence 6, Application US/08461146C

Patent No. 5981265

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: METHODS FOR REGULATING MEKK PROTEIN ACTIVITY

NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:

Lahive and Cockfield

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,146C

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/354,516

FILING DATE: 21-FEB-1995

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 15-APR-1993

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/323,460

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/11690

FILING DATE: 14-OCT-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/US94/04178

FILING DATE: 15-APR-1994

ATTORNEY/AGENT INFORMATION:

NAME: KARA, Catherine J.

REGISTRATION NUMBER: P41,106

REFERENCE/DOCKET NUMBER: CPI-004CN3

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 227-7400

TELEFAX: (617) 227-5941

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 626 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-461-146C-6

Query Match 90.3%; Score 28; DB 2; Length 626;

Best Local Similarity 83.3%; Pred. No. 2.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREDDV 6

Db 64 VREDDV 69

RESULT 16
US-08-461-145C-6
Sequence 6, Application US/08461145C

Patent No. 6074861

GENERAL INFORMATION:

APPLICANT: JOHNSON, GARY L.

TITLE OF INVENTION: NOVEL MEKK PROTEINS

NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

Lahive and Cockfield, LLP

STREET: 60 State Street

CITY: Boston

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/461,145C

FILING DATE: 5-JUNE-1995

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/049,254

FILING DATE: 11-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/323,460
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/11690
FILING DATE: 14-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/04178
FILING DATE: 15-APR-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/354,516
FILING DATE: 21-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: P-41,106
REFERENCE/DOCKET NUMBER: CPT-004CNI
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 626 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-145C-6

Query Match 90.3%; Score 28; DB 3; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||:
DB 64 VREYDV 69

RESULT 17
US-09-423-890-6
Sequence 6, Application US/09423890
Patent No. 6312934
GENERAL INFORMATION:
APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USSN 60/078,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USSN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 6
LENGTH: 626
TYPE: PRT
ORGANISM: Homo sapiens
US-09-423-890-6

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||:
DB 64 VREYDV 69

RESULT 18
US-09-423-890-12
Sequence 12, Application US/09423890
Patent No. 6312934
GENERAL INFORMATION:

APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES
FILE REFERENCE: CPT-085CPC
CURRENT APPLICATION NUMBER: US/09/423,890
PRIOR FILING DATE: 2000-03-06
PRIOR APPLICATION NUMBER: USSN 60/078,153
PRIOR FILING DATE: 1998-03-16
PRIOR APPLICATION NUMBER: USSN 60/099,165
PRIOR FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 12
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-09-423-890-12

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||:
DB 64 VREYDV 69

RESULT 19
US-08-628-829-10
Sequence 10, Application US/08628829A
Patent No. 6333170
GENERAL INFORMATION:
APPLICANT: Johnson, Gary L.
TITLE OF INVENTION: Method And Product For Regulating Cell Responsiveness To Extr
FILE REFERENCE: CPT-004DVC3
CURRENT APPLICATION NUMBER: US/08/628,829A
PRIOR FILING DATE: 1996-04-05
PRIOR APPLICATION NUMBER: 08/440,421
PRIOR FILING DATE: 1995-05-15
PRIOR APPLICATION NUMBER: 08/323,460
PRIOR FILING DATE: 1994-10-14
PRIOR APPLICATION NUMBER: 08/049,254
PRIOR FILING DATE: 1993-05-15
PRIOR APPLICATION NUMBER: 08/410,602
PRIOR FILING DATE: 1995-04-24
PRIOR APPLICATION NUMBER: 08/472,934
PRIOR FILING DATE: 1995-06-06
NUMBER OF SEQ ID NOS: 25
SOFTWARE: Patentln Ver. 2.0
SEQ ID NO 10
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-08-628-829-10

Query Match 90.3%; Score 28; DB 4; Length 626;
Best Local Similarity 83.3%; Pred. No. 2.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREYDL 6
|||||:
DB 64 VREYDV 69

RESULT 20
US-08-411-607A-2
Sequence 2, Application US/08411607A
Patent No. 5944102
GENERAL INFORMATION:
APPLICANT: HUDSON, PETER L
APPLICANT: ROSEN, CRAIG A
APPLICANT: HE, WEI WJ
TITLE OF INVENTION: PROSTATIC GROWTH FACTOR

NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
ADDRESSEE: STEWART & OLSTEIN
STREET: 6 BECKER FARM ROAD
CITY: ROSELAND
STATE: NJ
COUNTRY: US
ZIP: 07068
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/411,607A
FILING DATE: 11-Apr-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: FERRARO, GREGORY D
REGISTRATION NUMBER: 36,134
REFERENCE/DOCKET NUMBER: 325800-329
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 994-1700
TELEFAX: (201) 994-1744
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 295 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-411-607A-2

Query Match 87.1%; Score 27; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 46 RYEDL 50

RESULT 21
US-08-927-433-9
Sequence 9, Application US/08927433
Patent No. 6107476
GENERAL INFORMATION:
APPLICANT: Elander, Mark G.
APPLICANT: Huang, Shaoming
APPLICANT: Jackson, Michael A.
APPLICANT: Peterson, Per A.
TITLE OF INVENTION: PROTEIN GROWTH FACTOR FOR TREATING PROSTATE CANCER, AND RELATE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Johnson & Johnson
STREET: One J & J Plaza
CITY: New Brunswick
STATE: New Jersey
COUNTRY: USA
ZIP: 08933
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/927,433
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Morrison, Alan J.
REGISTRATION NUMBER: 37,399
REFERENCE/DOCKET NUMBER: ONT-849

TELECOMMUNICATION INFORMATION:
TELEPHONE: 732-524-3592
TELEFAX: 732-524-2808
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-927-433-9

Query Match 87.1%; Score 27; DB 3; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 22
US-08-775-882-4
Sequence 4, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: OH, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAKI, Mihoro
APPLICANT: KOBAYASHI, Midori
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
TITLE OF INVENTION: AGENT
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-4

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 23
US-08-775-882-6
Sequence 6, Application US/08775882
Patent No. 6180602
GENERAL INFORMATION:
APPLICANT: KATO, Seishi
APPLICANT: OH, Suwan
APPLICANT: SEKINE, Shingo
APPLICANT: SAKET, Miboro
APPLICANT: KOBAYASHI, Mido-ri
APPLICANT: YADA, Mika
APPLICANT: TSUJI, Tomoko
APPLICANT: OHMORI, Hitoshi
TITLE OF INVENTION: HUMAN NOVEL CDNA, TGF-BETA SUPERFAMILY
TITLE OF INVENTION: PROTEIN ENCODED THEREBY AND THE USE AS IMMUNOSUPPRESSIVE
TITLE OF INVENTION: AGENT
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,882
FILING DATE: 02-JAN-1997
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/390,207
FILING DATE: 16-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/379,441
FILING DATE: 03-FEB-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 5-61431
FILING DATE: 26-FEB-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-327619
FILING DATE: 13-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 4-208077
FILING DATE: 04-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: WEGNER, Harold C.
REGISTRATION NUMBER: 25,258
REFERENCE/DOCKET NUMBER: 53466/194

TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 308 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-775-882-6

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 24
US-09-276-600-6
Sequence 6, Application US/09276600
Patent No. 6465181
GENERAL INFORMATION:
APPLICANT: Patricia Billing-Medel
APPLICANT: Maurice Cohen
APPLICANT: Tracey L. Colpitts
APPLICANT: Julian Gordon
APPLICANT: Edward N. Granados
APPLICANT: John C. Russell
APPLICANT: Stephen D. Stroupe
TITLE OF INVENTION: Reagents and Methods Useful for
TITLE OF INVENTION: Detecting Disease of the Prostate
FILE REFERENCE: 6397.US.01
CURRENT APPLICATION NUMBER: US/09/276,600
CURRENT FILING DATE: 1999-03-25
NUMBER OF SEQ ID NOS: 11
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 6
LENGTH: 308
TYPE: PRT
ORGANISM: Homo sapiens
US-09-276-600-6

Query Match 87.1%; Score 27; DB 4; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 59 RYEDL 63

RESULT 25
US-09-067-351-2
Sequence 2, Application US/09067351
Patent No. 5994081
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baughn, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/067,351
FILING DATE: Herewith
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 2029060
US-09-067-351-2

Query Match 87.1%; Score 27; DB 2; Length 551;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 237 VRYED 241

RESULT 26
US-09-360-490-2
Sequence 2, Application US/09360490
Patent No. 6221843
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Baugim, Mariah
TITLE OF INVENTION: HUMAN KERATINS
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/360,490
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/067,351
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: CERRONE, MICHAEL C.
REGISTRATION NUMBER: 39,132
REFERENCE/DOCKET NUMBER: PF-0511 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 551 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: KERANOT02
CLONE: 2029060
US-09-360-490-2

Query Match 87.1%; Score 27; DB 4; Length 551;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 237 VRYED 241

RESULT 27
US-09-041-236-4
Sequence 4, Application US/09041236
Patent No. 6225285
GENERAL INFORMATION:
APPLICANT: Luo, Yuling
APPLICANT: Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,236
FILING DATE: March 11, 1998
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-236-4

Query Match 87.1%; Score 27; DB 4; Length 606;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
Db 415 VRYED 419

RESULT 28
US-09-412-554A-2
Sequence 2, Application US/09412554A
Patent No. 6355788

GENERAL INFORMATION:
APPLICANT: Conklin, Darrell
APPLICANT: Ellsworth, Jeff L.
TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTX2
FILE REFERENCE: 98-50
CURRENT APPLICATION NUMBER: US/09/412,554A
CURRENT FILING DATE: 1999-10-05
NUMBER OF SEQ ID NOS: 15
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 983
TYPE: PRT
ORGANISM: Homo sapiens
US-09-412-554A-2

Query Match 87.1%; Score 27; DB 4; Length 983;
Best Local Similarity 100.0%; Pred. No. 6.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 539 VRYED 543

RESULT 29
US-09-396-651B-2
Sequence 2, Application US/09396651B
Patent No. 6225076
GENERAL INFORMATION:
APPLICANT: Darst, Seth A
APPLICANT: Zhang, Gongyi
APPLICANT: Campbell, Elizabeth
APPLICANT: Minakin, Leonid
APPLICANT: Severinov, Konstantin
TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
FILE REFERENCE: 600-1-258
CURRENT APPLICATION NUMBER: US/09/396,651B
CURRENT FILING DATE: 1999-09-15
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2
LENGTH: 1119
TYPE: PRT
ORGANISM: Thermus aquaticus
FEATURE:
NAME/KEY: SITE
LOCATION: (695)..(696)
OTHER INFORMATION: Any amino acids can be at these two positions.
US-09-396-651B-2

Query Match 87.1%; Score 27; DB 4; Length 1119;
Best Local Similarity 100.0%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 613 VRYED 617

RESULT 30
US-08-117-083-60
Sequence 60, Application US/08117083
Patent No. 5719054
GENERAL INFORMATION:
APPLICANT: Bourisnelli, Michael E.
APPLICANT: Inglis, Stephen C.
APPLICANT: Munro, Alan J.
TITLE OF INVENTION: Recombinant Virus Vectors Encoding Human
TITLE OF INVENTION: Papilloma Virus Proteins
NUMBER OF SEQUENCES: 70
CORRESPONDENCE ADDRESS:
ADDRESSEE: Walter H. Dreger

STREET: 4 Embarcadero Center, Suite 3400
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,083
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Dreger, Walter H.
REGISTRATION NUMBER: 24,190
REFERENCE/DOCKET NUMBER: A-58783
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-781-1989
TELEFAX: 415-398-3249
TELEX: 910 277299

INFORMATION FOR SEQ ID NO: 60:
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
NAME/KEY: Protein
LOCATION: 1..416

OTHER INFORMATION: /note- "Xaa refers to stop codon in
the open reading frame."
US-08-117-083-60

Query Match 83.9%; Score 26; DB 1; Length 416;
Best Local Similarity 66.7%; Pred. No. 3.9e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYEDL 6
DB 341 IYEDL 346

RESULT 31
US-08-588-963-20
Sequence 20, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:

APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P. O. Box 4433
CITY: Houston
STATE: TX

COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Pussey, Shelley P. M.
REGISTRATION NUMBER: 39,458

REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: L/A
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-20

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 154 VRHEDL 159

RESULT 32

US-08-588-983-22
Sequence 22, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-22

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 154 VRHEDL 159

RESULT 33

US-08-588-976-20
Sequence 20, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-20

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 VREEDL 6
DB 154 VRHEDL 159

RESULT 34
US-08-588-976-22
Sequence 22, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 465 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-22

Query Match 83.9%; Score 26; DB 2; Length 465;
Best Local Similarity 83.3%; Pred. No. 4.4e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 154 VRHEDL 159

RESULT 35
US-08-588-983-9
Sequence 9, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-9

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||

Db 608 VRHEDL 613

RESULT 36
US-08-588-983-12
Sequence 12, Application US/08588983
Patent No. 5854067
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions
TITLE OF INVENTION: for Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,983
FILING DATE: Concurrently herewith
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:424/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-983-12

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
|||||
Db 608 VRHEDL 613

RESULT 37
US-08-588-976-9
Sequence 9, Application US/08588976
Patent No. 5891717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-9

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRXEDL 6
DB 608 VRHEDL 613

RESULT 38
US-08-588-976-12
Sequence 12, Application US/08588976
Patent No. 5691717
GENERAL INFORMATION:
APPLICANT: Christopher B. Newgard, et al.
TITLE OF INVENTION: Methods and Compositions for
TITLE OF INVENTION: Inhibiting Hexokinase
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Arnold, White & Durkee
STREET: P.O. Box 4433
CITY: Houston
STATE: TX
COUNTRY: US
ZIP: 77210
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/588,976
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Fussey, Shelley P.M.
REGISTRATION NUMBER: 39,458
REFERENCE/DOCKET NUMBER: UTSD:481/FUS
TELECOMMUNICATION INFORMATION:
TELEPHONE: (512) 418-3000
TELEFAX: (512) 474-7577
TELEX: n/a
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 919 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-588-976-12

Query Match 83.9%; Score 26; DB 2; Length 919;
Best Local Similarity 83.3%; Pred. No. 9.5e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRXEDL 6
DB 608 VRHEDL 613

RESULT 39
US-08-040-751-3
Sequence 3, Application US/08040751
Patent No. 5407825
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M
APPLICANT: Sick, August J
TITLE OF INVENTION: No. 5407825el Bacillus thuringiensis isolates
TITLE OF INVENTION: active against Lepidopteran Pests and Genes Encoding No. 54
TITLE OF INVENTION: Lepidopteran-active Toxins
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESSEE: DAVID R. SALIMANCHIK
STREET: 2421 N.W. 41st STREET, SUITE A-1
CITY: GAINESVILLE
STATE: FL
COUNTRY: USA
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/040,751
FILING DATE: 19930329
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: SALIMANCHIK, DAVID R.
REFERENCE/DOCKET NUMBER: MA39.C1.D3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 904-375-8100
TELEFAX: 904-372-5800
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: AMINO ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: Lambdaagem - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-040-751-3

Query Match 83.9%; Score 26; DB 1; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VRXED 5
DB 1087 IRYED 1091

RESULT 40
US-08-291-368-2
Sequence 2, Application US/08291368
Patent No. 5686069
GENERAL INFORMATION:

APPLICANT: Payne, Jewel M.
APPLICANT: Sick, August J.
TITLE OF INVENTION: No. 568609e1 Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/291,368
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA50.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: Lambdagem - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-291-368-2
Query Match 83.9%; Score 26; DB 1; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VARIED 5
DB 1087 IRYED 1091
RESULT 41
US-08-962-190-2
Sequence 2, Application US/08962190
Patent No. 5985267
GENERAL INFORMATION:
APPLICANT: Payne, Jewel M.
APPLICANT: Sick, August J.
TITLE OF INVENTION: No. 5985267e1 Bacillus thuringiensis Isolates
TITLE OF INVENTION: Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville

STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/962,190
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/291,368
FILING DATE:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MA50.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ. ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: Lambdagem - 11 (tm) Library of August Sick
CLONE: 81A2
US-08-962-190-2
Query Match 83.9%; Score 26; DB 2; Length 1174;
Best Local Similarity 80.0%; Pred. No. 1.3e+03;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 VARIED 5
DB 1087 IRYED 1091
RESULT 42
PCT-US95-10310-2
Sequence 2, Application PC/TUS9510310
GENERAL INFORMATION:
APPLICANT: MYCOGEN CORPORATION
APPLICANT: STREET ADDRESS: 5501 Oberlin Drive
APPLICANT: CITY: San Diego
APPLICANT: STATE/PROVINCE: California
APPLICANT: COUNTRY: US
APPLICANT: POSTAL CODE/ZIP: 92121
APPLICANT: PHONE NUMBER: (619) 453-8030
APPLICANT: FAX NUMBER: (619) 453-6991
TITLE OF INVENTION: Protein Toxins Active Against Lepidopteran Pests
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSER: Saliwanchik & Saliwanchik
STREET: 2421 N.W. 41st Street, Suite A-1
CITY: Gainesville
STATE: FL
COUNTRY: US
ZIP: 32606
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/10310
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/597,607
FILING DATE: 15-OCT-90
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Saliwanchik, David R.
REGISTRATION NUMBER: 31,794
REFERENCE/DOCKET NUMBER: MAS0.C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (904)375-8100
TELEFAX: (904)372-5800
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1174 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHEICAL: YES
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Bacillus thuringiensis
STRAIN: aizawai
INDIVIDUAL ISOLATE: PS81A2
IMMEDIATE SOURCE:
LIBRARY: lambdagem - 11 (tm) Library of August Sick
CLONE: 81A2
PCT-US95-10310-2

Query Match
Best Local Similarity 83.9%; Score 26; DB 5; Length 1174;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 5
:|||||
Db 1087 IRYED 1091

RESULT 43
5164180-4
; Patent No. 5164180
; APPLICANT: Payne, Jewel; Sick, August J.
; TITLE OF INVENTION: BACILLUS THURINGIENSIS ISOLATES ACTIVE
; AGAINST LEPIDOPTERAN PESTS
; NUMBER OF SEQUENCES: 6
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/451,389
; FILING DATE: 14-DEC-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 353,860
; FILING DATE: 18-MAY-1989
; SEQ ID NO: 4:
; LENGTH: 1174
5164180-4

Query Match
Best Local Similarity 83.9%; Score 26; DB 6; Length 1174;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 5
:|||||
Db 1087 IRYED 1091

RESULT 44

US-09-172-422-1
; Sequence 1, Application US/09172422A
; Patent No. 6300485
; GENERAL INFORMATION:
; APPLICANT: Adams, Arwen E.
; APPLICANT: Chiu, Choi Ying
; APPLICANT: Duhl, David
; APPLICANT: Gorman, Susan W.
; APPLICANT: Leng, Song
; APPLICANT: Sheffield, Val
; APPLICANT: Welch, Juliet
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED
; TITLE OF INVENTION: CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,
; FILE REFERENCE: 200130.442
; CURRENT APPLICATION NUMBER: US/09/172,422A
; CURRENT FILING DATE: 1998-10-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 2548
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-172-422-1

Query Match
Best Local Similarity 83.9%; Score 26; DB 4; Length 2548;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VYED 6
:|||||
Db 1920 IRYKDL 1925

RESULT 45
US-09-370-253-14
; Sequence 14, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-370-253-14

Query Match
Best Local Similarity 80.6%; Score 25; DB 4; Length 103;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 VYED 6
:|||||
Db 86 IRYIDL 91

RESULT 46
US-09-352-990-20
; Sequence 20, Application US/09352990
; Patent No. 6255090
; GENERAL INFORMATION:
; APPLICANT: Famodu, Layo O.
; APPLICANT: Orozco, Buddy
; APPLICANT: Rafalski, Antoni

FILE REFERENCE: BB-1191
CURRENT APPLICATION NUMBER: US/09/352,990
CURRENT FILING DATE: 1999-07-14
EARLIER APPLICATION NUMBER: 60/092,866
EARLIER FILING DATE: JULY 15, 1998
NUMBER OF SEQ ID NOS: 29
SOFTWARE: Microsoft Office 97
SEQ ID NO 20
LENGTH: 171
TYPE: PRT
ORGANISM: Trilicium aestivum
US-09-352-990-20

Query Match 80.6%; Score 25; DB 4; Length 171;
Best Local Similarity 66.7%; Pred. No. 2.4e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 128 VRYEEL 133

RESULT 47
US-08-698-805-4
Sequence 4, Application US/08698805
Patent No. 5869288
GENERAL INFORMATION:
APPLICANT: Chapman, L. Karla
APPLICANT: Attitude, L. Karla
TITLE OF INVENTION: Molecular Cloning of Cockroach
TITLE OF INVENTION: Allergens, Amino Acid and Nucleotide Sequences Therefore,
TITLE OF INVENTION: and Recombinant Expression Thereof
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Oblon, Spivak, McClelland, Maiter & Neustadt,
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/698,805
FILING DATE: 16-AUG-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,510
FILING DATE: 18-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Kelber, Steven B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 49-203-27
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-413-3000
TELEFAX: 703-413-2220
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-698-805-4

Query Match 80.6%; Score 25; DB 2; Length 182;
Best Local Similarity 66.7%; Pred. No. 2.6e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 172 LRYEDL 177

RESULT 48
US-09-210-748A-2
Sequence 2, Application US/09210748A
Patent No. 635156
GENERAL INFORMATION:
APPLICANT: Hermeking, Helko
APPLICANT: Vogelstein, Bert
APPLICANT: Kinzler, Kenneth
TITLE OF INVENTION: 14-3-3 SIGMA ARREST THE CELL CYCLE
FILE REFERENCE: 1107.77810
CURRENT APPLICATION NUMBER: US/09/210,748A
CURRENT FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/069,416
PRIOR FILING DATE: 1997-12-18
NUMBER OF SEQ ID NOS: 18
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 248
TYPE: PRT
ORGANISM: Homo sapiens
US-09-210-748A-2

Query Match 80.6%; Score 25; DB 4; Length 248;
Best Local Similarity 80.0%; Pred. No. 3.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 18 RYEDM 22

RESULT 49
US-07-887-072B-2
Sequence 2, Application US/07887072B
Patent No. 5424191
GENERAL INFORMATION:
APPLICANT: Prasad Ph.D., Gaddamanugu L.
APPLICANT: Cooper M.D., Herbert L.
TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
TITLE OF INVENTION: MARKER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 W. Broadway, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/887,072B
FILING DATE: 20-MAY-1992
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kirkpatrick Ph.D., Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH021,021A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid

Search completed: February 20, 2003, 13:35:10
Job time: 16.8571 secs

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-07-887-072B-2

Query Match 80.6%; Score 25; DB 1; Length 249;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
Db 18 RYEDM 22

RESULT 50
US-08-466-444-2
Sequence 2, Application US/08466444
Patent No. 5776676
GENERAL INFORMATION:
APPLICANT: Prasad Ph.D., Gaddamanugu L.
APPLICANT: Cooper M.D., Heider L.
TITLE OF INVENTION: EPITHELIAL CELL SPECIFIC DIFFERENTIATION
TITLE OF INVENTION: MARKER
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe, Martens, Olson & Bear
STREET: 501 W. Broadway, Suite 1700
CITY: San Diego
STATE: California
COUNTRY: USA
ZIP: 92101
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/466,444
FILING DATE: 06-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/887,072
FILING DATE: 20-MAY-1992
ATTORNEY/AGENT INFORMATION:
NAME: Kirpatrick Ph.D., Anita M.
REGISTRATION NUMBER: 32,617
REFERENCE/DOCKET NUMBER: NIH021,021A
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 249 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
US-08-466-444-2

Query Match 80.6%; Score 25; DB 1; Length 249;
Best Local Similarity 80.0%; Pred. No. 3.7e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RYEDL 6
Db 18 RYEDM 22

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 20, 2003, 13:33:52 ; Search time 2.74286 Seconds
(without alignments)
55.888 Million cell updates/sec

Title: US-09-816-825-9
Perfect score: 31
Sequence: 1 VRYEDL 6

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 140259 seqs, 2554876 residues

Total number of hits satisfying chosen parameters: 140259

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/PCCTUS_PUBCOMB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	31	100.0	6	10	US-09-816-825-9
2	31	100.0	169	10	US-09-927-602-9
3	31	100.0	171	10	US-09-927-602-8
4	31	100.0	179	10	US-09-927-602-10
5	31	100.0	386	9	US-10-007-262-1
6	31	100.0	386	10	US-09-816-825-2
7	31	100.0	418	10	US-09-927-602-5
8	31	100.0	484	9	US-10-212-933-2
9	31	100.0	484	9	US-10-212-933-4
10	31	100.0	531	10	US-09-833-790-255
11	28	90.3	626	12	US-09-927-602-11
12	28	90.3	626	12	US-10-000-864-6
13	28	90.3	626	12	US-10-000-864-12
14	27	87.1	65	10	US-09-864-761-35570
15	27	87.1	65	10	US-09-864-761-49114
16	27	87.1	157	9	US-09-738-626-3600
17	27	87.1	169	10	US-09-927-602-6
18	27	87.1	169	10	US-09-927-602-7
19	27	87.1	295	10	US-09-361-741-2

20	27	87.1	308	9	US-09-981-353-188	Sequence 188, App
21	27	87.1	308	10	US-09-276-600-6	Sequence 6, Appli
22	27	87.1	316	9	US-09-966-546-14	Sequence 14, Appl
23	27	87.1	316	9	US-09-966-545-14	Sequence 14, Appl
24	27	87.1	316	9	US-09-965-212-14	Sequence 14, Appl
25	27	87.1	390	10	US-09-927-602-4	Sequence 4, Appli
26	27	87.1	395	10	US-09-927-602-2	Sequence 2, Appli
27	27	87.1	395	10	US-09-927-602-3	Sequence 3, Appli
28	27	87.1	396	10	US-09-815-242-12541	Sequence 12541, A
29	27	87.1	396	10	US-09-815-242-12900	Sequence 12900, A
30	27	87.1	439	10	US-09-741-669-300	Sequence 300, App
31	27	87.1	439	10	US-09-815-242-10284	Sequence 10284, A
32	27	87.1	604	10	US-09-907-509-2	Sequence 2, Appli
33	27	87.1	606	9	US-09-771-467C-4	Sequence 4, Appli
34	27	87.1	693	9	US-09-808-602-59	Sequence 59, Appl
35	27	87.1	773	9	US-09-808-602-60	Sequence 60, Appl
36	27	87.1	815	9	US-09-808-602-4	Sequence 4, Appli
37	27	87.1	815	9	US-09-966-546-18	Sequence 18, Appl
38	27	87.1	815	9	US-09-966-545-18	Sequence 18, Appl
39	27	87.1	815	9	US-09-965-212-18	Sequence 18, Appl
40	27	87.1	842	9	US-09-966-546-16	Sequence 16, Appl
41	27	87.1	842	9	US-09-966-545-16	Sequence 16, Appl
42	27	87.1	842	9	US-09-965-212-16	Sequence 16, Appl
43	27	87.1	850	9	US-09-808-602-58	Sequence 58, Appl
44	27	87.1	850	9	US-09-808-602-73	Sequence 73, Appl
45	27	87.1	983	9	US-10-013-136-2	Sequence 2, Appli
46	27	87.1	1119	10	US-09-782-714-2	Sequence 2, Appli
47	26	83.9	45	10	US-09-864-761-34609	Sequence 34609, A
48	26	83.9	144	10	US-09-764-878-127	Sequence 127, App
49	26	83.9	144	10	US-09-764-860-461	Sequence 461, App
50	26	83.9	135	9	US-09-764-868-919	Sequence 919, App
51	26	83.9	332	9	US-09-764-868-1225	Sequence 1225, App
52	26	83.9	407	9	US-09-738-626-5541	Sequence 5541, Ap
53	26	83.9	904	9	US-09-964-899-49	Sequence 49, Appl
54	26	83.9	2548	10	US-09-851-682A-1	Sequence 1, Appli
55	25	80.6	55	10	US-09-925-299-1150	Sequence 1150, App
56	25	80.6	146	10	US-09-747-835A-16	Sequence 16, Appl
57	25	80.6	195	10	US-09-815-028-8	Sequence 8, Appli
58	25	80.6	248	10	US-09-939-581A-2	Sequence 2, Appli
59	25	80.6	371	10	US-09-815-242-10916	Sequence 10916, A
60	25	80.6	373	9	US-09-771-161A-201	Sequence 201, App
61	25	80.6	383	10	US-09-771-161A-110	Sequence 110, App
62	25	80.6	414	10	US-09-771-161A-111	Sequence 111, App
63	25	80.6	488	10	US-09-815-242-11444	Sequence 11444, A
64	25	80.6	498	9	US-09-893-519A-33	Sequence 33, Appli
65	25	80.6	599	10	US-09-789-561-120	Sequence 120, App
66	25	80.6	626	10	US-09-771-161A-201	Sequence 201, App
67	25	80.6	832	10	US-09-771-161A-202	Sequence 202, App
68	25	80.6	928	10	US-09-919-585-21	Sequence 21, Appl
69	25	80.6	928	10	US-09-801-368-282	Sequence 282, App
70	24	77.4	30	10	US-09-864-761-45818	Sequence 45818, A
71	24	77.4	67	10	US-09-864-761-39410	Sequence 39410, A
72	24	77.4	79	10	US-09-864-761-45955	Sequence 45955, A
73	24	77.4	93	10	US-09-939-980-466	Sequence 466, App
74	24	77.4	98	10	US-09-840-459-40	Sequence 40, Appl
75	24	77.4	139	9	US-09-796-692-1084	Sequence 1084, Ap
76	24	77.4	139	9	US-09-796-692-1558	Sequence 1558, Ap
77	24	77.4	139	9	US-09-796-692-1955	Sequence 1955, Ap
78	24	77.4	139	9	US-09-796-692-2194	Sequence 2194, Ap
79	24	77.4	143	9	US-09-934-465-12	Sequence 12, Appl
80	24	77.4	148	10	US-09-925-299-1164	Sequence 1164, Ap
81	24	77.4	186	10	US-09-802-116-2	Sequence 2, Appli
82	24	77.4	187	10	US-09-802-116-2	Sequence 2, Appli
83	24	77.4	209	10	US-09-925-299-1038	Sequence 1038, Ap
84	24	77.4	233	9	US-09-738-626-5648	Sequence 5648, Ap
85	24	77.4	234	10	US-09-813-229-16	Sequence 16, Appl
86	24	77.4	234	10	US-09-921-667-2	Sequence 2, Appli
87	24	77.4	239	10	US-09-795-440A-2	Sequence 2, Appli
88	24	77.4	239	10	US-09-795-440A-4	Sequence 4, Appli
89	24	77.4	239	12	US-10-060-857-1	Sequence 1, Appli
90	24	77.4	240	10	US-09-819-817-7	Sequence 7, Appli
91	24	77.4	271	10	US-09-872-733-7	Sequence 7, Appli
92	24	77.4	271	10	US-09-872-733-17	Sequence 17, Appli

239	23	74.2	313	10	US-09-815-242-10920	Sequence 10920, A	312	23	74.2	858	10	US-09-986-224-8	Sequence 8, Appl1
240	23	74.2	345	10	US-09-815-242-13259	Sequence 13259, A	313	23	74.2	876	10	US-09-815-242-13003	Sequence 13003, A
241	23	74.2	357	10	US-09-815-242-10826	Sequence 10826, A	314	23	74.2	879	10	US-09-815-242-5312	Sequence 5312, Ap
242	23	74.2	357	10	US-09-815-242-11389	Sequence 11389, A	315	23	74.2	883	10	US-09-815-242-13382	Sequence 13382, A
243	23	74.2	359	10	US-09-815-242-13336	Sequence 13336, A	316	23	74.2	883	10	US-09-815-242-13364	Sequence 13364, A
244	23	74.2	359	10	US-09-799-777-45	Sequence 45, Appl	317	23	74.2	883	10	US-09-815-242-13868	Sequence 6108, Ap
245	23	74.2	360	10	US-09-815-242-5195	Sequence 5195, Ap	318	23	74.2	900	9	US-09-738-626-6108	Sequence 6108, Ap
246	23	74.2	361	10	US-09-815-242-11251	Sequence 11251, A	319	23	74.2	904	10	US-09-894-998-18	Sequence 18, Appl
247	23	74.2	371	9	US-10-137-731-2	Sequence 15, Appl	320	23	74.2	951	10	US-09-817-514A-6	Sequence 6, Appl
248	23	74.2	373	10	US-09-828-302-15	Sequence 4866, Ap	321	23	74.2	951	10	US-09-815-242-10465	Sequence 10465, A
249	23	74.2	382	9	US-09-738-626-4866	Sequence 36, Appl	322	23	74.2	994	9	US-10-066-571-24	Sequence 24, Appl
250	23	74.2	390	10	US-09-764-848-36	Sequence 64, Appl	323	23	74.2	1037	9	US-09-728-721-55	Sequence 55, Appl
251	23	74.2	396	9	US-09-971-536-64	Sequence 12, Appl	324	23	74.2	1127	10	US-09-815-242-5373	Sequence 5373, Ap
252	23	74.2	415	10	US-09-971-361-12	Sequence 18, Appl	325	23	74.2	1148	10	US-09-815-242-11895	Sequence 11895, A
253	23	74.2	422	10	US-09-779-307-18	Sequence 20, Appl	326	23	74.2	1153	10	US-09-992-056-2	Sequence 2, Appl
254	23	74.2	422	10	US-09-779-307-20	Sequence 3778, Ap	327	23	74.2	1158	10	US-09-815-242-12522	Sequence 12522, A
255	23	74.2	423	9	US-09-738-626-3778	Sequence 16, Appl	328	23	74.2	1242	10	US-09-904-065-12	Sequence 13, Appl
256	23	74.2	424	10	US-09-733-524-16	Sequence 6, Appl	329	23	74.2	1242	10	US-09-904-065-13	Sequence 3, Appl
257	23	74.2	425	9	US-10-120-319-6	Sequence 21, Appl	330	23	74.2	1332	9	US-10-041-856-3	Sequence 3, Appl
258	23	74.2	428	10	US-09-779-307-21	Sequence 19, Appl	331	23	74.2	1332	9	US-09-738-626-4048	Sequence 4048, Ap
259	23	74.2	430	10	US-09-779-307-19	Sequence 2, Appl	332	23	74.2	1332	9	US-09-864-761-46862	Sequence 46862, A
260	23	74.2	432	10	US-09-815-242-5096	Sequence 17, Appl	333	23	74.2	1797	10	US-09-881-752A-308	Sequence 308, App
261	23	74.2	454	9	US-10-060-425-17	Sequence 8, Appl	334	23	74.2	1958	12	US-10-028-446-4	Sequence 4, Appl
262	23	74.2	454	9	US-10-120-319-8	Sequence 17, Appl	335	23	74.2	2053	9	US-10-017-216-2	Sequence 2, Appl1
263	23	74.2	454	10	US-09-827-040-4	Sequence 4, Appl	336	23	74.2	2054	12	US-10-028-446-2	Sequence 2, Appl1
264	23	74.2	454	10	US-09-733-524-18	Sequence 18, Appl	337	23	74.2	2055	9	US-10-017-216-4	Sequence 4, Appl1
265	23	74.2	456	10	US-09-815-242-10870	Sequence 10870, A	338	23	74.2	2890	10	US-09-815-242-11410	Sequence 11410, A
266	23	74.2	464	9	US-10-120-319-1	Sequence 1, Appl1	339	23	74.2	2890	10	US-09-815-242-11571	Sequence 11571, A
267	23	74.2	473	10	US-09-815-242-11235	Sequence 11235, A	340	23	74.2	2890	10	US-09-870-759-128	Sequence 128, App
268	23	74.2	476	9	US-10-120-319-5	Sequence 5, Appl1	341	22	71.0	26	10	US-09-749-334A-5	Sequence 5, Appl
269	23	74.2	476	10	US-09-733-524-15	Sequence 15, Appl	342	22	71.0	41	10	US-09-864-761-37339	Sequence 37339, A
270	23	74.2	476	10	US-09-779-307-16	Sequence 4, Appl	343	22	71.0	46	10	US-09-864-761-42065	Sequence 42065, A
271	23	74.2	477	10	US-09-779-307-4	Sequence 4, Appl	344	22	71.0	46	10	US-09-864-761-42065	Sequence 42065, A
272	23	74.2	478	9	US-10-120-319-7	Sequence 7, Appl1	345	22	71.0	46	10	US-09-864-761-42065	Sequence 42065, A
273	23	74.2	479	10	US-09-733-524-17	Sequence 17, Appl	346	22	71.0	50	10	US-09-938-315-50	Sequence 31753, Ap
274	23	74.2	480	10	US-09-802-472B-4	Sequence 4, Appl1	347	22	71.0	52	10	US-09-864-761-42113	Sequence 42113, A
275	23	74.2	482	10	US-09-779-307-17	Sequence 17, Appl	348	22	71.0	62	9	US-09-796-652-1724	Sequence 1724, Ap
276	23	74.2	483	9	US-09-974-298-41	Sequence 41, Appl	349	22	71.0	62	10	US-09-864-761-39904	Sequence 39904, A
277	23	74.2	483	10	US-09-981-353-158	Sequence 158, App	350	22	71.0	70	10	US-09-925-999-1218	Sequence 1218, Ap
278	23	74.2	483	10	US-09-919-497-79	Sequence 79, Appl	351	22	71.0	78	9	US-09-940-240-8	Sequence 240-8, Ap
279	23	74.2	485	9	US-09-733-524-2	Sequence 2, Appl1	352	22	71.0	79	9	US-09-738-626-4440	Sequence 4440, Ap
280	23	74.2	486	10	US-10-120-319-2	Sequence 2, Appl1	353	22	71.0	89	9	US-09-738-626-5666	Sequence 5666, Ap
281	23	74.2	490	9	US-09-895-913A-6	Sequence 6, Appl1	354	22	71.0	93	10	US-09-764-860-518	Sequence 518, App
282	23	74.2	492	9	US-09-796-753-42	Sequence 42, Appl	355	22	71.0	94	9	US-09-940-240-21	Sequence 21, Appl
283	23	74.2	494	9	US-10-238-709-4	Sequence 4, Appl1	356	22	71.0	99	9	US-09-940-240-2	Sequence 2, Appl1
284	23	74.2	494	10	US-09-804-471A-4	Sequence 4, Appl1	357	22	71.0	99	9	US-09-978-285A-370	Sequence 370, App
285	23	74.2	495	10	US-09-815-242-10315	Sequence 10315, A	358	22	71.0	111	9	US-09-978-285A-370	Sequence 370, App
286	23	74.2	497	10	US-10-238-709-2	Sequence 2, Appl1	359	22	71.0	111	9	US-09-978-285A-370	Sequence 370, App
287	23	74.2	497	10	US-09-804-471A-2	Sequence 2, Appl1	360	22	71.0	111	9	US-09-978-285A-370	Sequence 370, App
288	23	74.2	501	9	US-09-738-626-6834	Sequence 6834, Ap	361	22	71.0	111	9	US-09-978-285A-370	Sequence 370, App
289	23	74.2	501	10	US-09-733-524-2	Sequence 1, Appl1	362	22	71.0	111	9	US-09-978-285A-370	Sequence 370, App
290	23	74.2	502	10	US-09-895-072-13	Sequence 13, Appl	363	22	71.0	111	10	US-09-978-285A-370	Sequence 370, App
291	23	74.2	502	10	US-09-986-552-13	Sequence 5577, Ap	364	22	71.0	111	10	US-09-978-285A-370	Sequence 370, App
292	23	74.2	528	9	US-09-738-626-5577	Sequence 5577, Ap	365	22	71.0	117	10	US-09-816-920-2	Sequence 2, Appl1
293	23	74.2	551	10	US-09-802-472B-21	Sequence 21, Appl	366	22	71.0	125	9	US-09-815-242-5266	Sequence 5266, Ap
294	23	74.2	560	10	US-09-864-761-37938	Sequence 37938, A	367	22	71.0	133	10	US-10-014-717-12	Sequence 12, Appl
295	23	74.2	572	10	US-09-864-761-44546	Sequence 34546, A	368	22	71.0	144	10	US-09-852-976-2	Sequence 2, Appl1
296	23	74.2	590	9	US-09-331-631A-8	Sequence 8, Appl1	369	22	71.0	149	10	US-09-734-569-30	Sequence 30, Appl
297	23	74.2	607	9	US-09-344-882-18	Sequence 16, Appl	370	22	71.0	149	10	US-09-764-868-1022	Sequence 1022, Ap
298	23	74.2	619	10	US-09-738-626-5531	Sequence 5531, Ap	371	22	71.0	151	10	US-09-804-969-8	Sequence 8, Appl1
299	23	74.2	729	10	US-09-815-242-5520	Sequence 5520, Ap	372	22	71.0	155	9	US-09-815-242-12807	Sequence 12807, A
300	23	74.2	750	10	US-09-815-242-10783	Sequence 10783, A	373	22	71.0	158	9	US-10-014-717-12	Sequence 12, Appl
301	23	74.2	762	10	US-09-391-340-10	Sequence 10, Appl	374	22	71.0	158	9	US-09-738-626-6468	Sequence 6468, Ap
302	23	74.2	762	10	US-09-948-369-10	Sequence 10, Appl	375	22	71.0	159	10	US-09-749-334A-30	Sequence 30, Appl
303	23	74.2	782	10	US-09-815-242-12367	Sequence 12367, A	376	22	71.0	162	10	US-09-925-302-593	Sequence 593, Appl
304	23	74.2	782	10	US-09-815-242-12984	Sequence 12984, A	377	22	71.0	176	10	US-09-738-626-6015	Sequence 6015, Ap
305	23	74.2	798	10	US-09-978-248-2	Sequence 2, Appl1	378	22	71.0	186	10	US-09-815-242-11338	Sequence 11338, A
306	23	74.2	798	10	US-09-765-519-2	Sequence 2, Appl1	379	22	71.0	187	10	US-09-917-265-102	Sequence 102, Ap
307	23	74.2	810	10	US-09-815-242-12636	Sequence 12636, A	380	22	71.0	187	10	US-09-925-297-490	Sequence 490, App
308	23	74.2	820	10	US-09-815-242-5882	Sequence 5882, Ap	381	22	71.0	193	9	US-09-738-626-3784	Sequence 3784, Appl
309	23	74.2	825	10	US-09-815-242-13079	Sequence 13079, A	382	22	71.0	193	10	US-09-941-936A-2	Sequence 2, Appl1
310	23	74.2	844	10	US-09-815-242-10039	Sequence 10039, A	383	22	71.0	194	9	US-10-025-380-1126	Sequence 1126, Ap
311	23	74.2	858	10	US-09-848-035-8	Sequence 8, Appl1	384	22	71.0	194	9	US-09-738-626-6465	Sequence 6465, Ap

385	22	71.0	195	9	US-10-074-956-23	Sequence 23, Appl	458	22	71.0	264	9	US-09-944-160-46	Sequence 49, Appl
386	22	71.0	197	10	US-09-917-265-36	Sequence 36, Appl	459	22	71.0	265	9	US-09-738-626-4442	Sequence 4442, Ap
387	22	71.0	197	10	US-09-917-265-50	Sequence 50, Appl	460	22	71.0	265	10	US-09-815-242-13460	Sequence 13460, A
388	22	71.0	197	12	US-10-079-616-4	Sequence 4, Appl1	461	22	71.0	266	10	US-09-815-242-13665	Sequence 13665, A
389	22	71.0	202	10	US-09-925-299-974	Sequence 974, App	462	22	71.0	266	10	US-09-925-197-2	Sequence 2, Appl1
390	22	71.0	203	9	US-10-174-590-270	Sequence 270, App	463	22	71.0	267	10	US-09-825-555-2	Sequence 2, Appl1
391	22	71.0	203	9	US-10-176-758-270	Sequence 270, App	464	22	71.0	268	10	US-09-925-300-1461	Sequence 241, App
392	22	71.0	203	9	US-10-175-737-270	Sequence 270, App	465	22	71.0	269	9	US-10-108-605-241	Sequence 4919, Ap
393	22	71.0	203	9	US-10-173-706-270	Sequence 270, App	466	22	71.0	274	10	US-09-815-242-4919	Sequence 5621, Ap
394	22	71.0	203	9	US-10-175-738-270	Sequence 270, App	467	22	71.0	275	9	US-09-815-242-5621	Sequence 144, App
395	22	71.0	203	9	US-10-175-752-270	Sequence 270, App	468	22	71.0	276	10	US-09-815-242-10555	Sequence 10555, A
396	22	71.0	203	9	US-10-176-482-270	Sequence 270, App	469	22	71.0	282	10	US-09-815-242-11733	Sequence 11733, A
397	22	71.0	203	9	US-10-176-757-270	Sequence 270, App	470	22	71.0	287	10	US-09-734-569-152	Sequence 152, App
398	22	71.0	203	9	US-10-176-913-270	Sequence 270, App	471	22	71.0	288	10	US-09-755-016-6	Sequence 6, Appl1
399	22	71.0	203	9	US-10-180-552-270	Sequence 270, App	472	22	71.0	290	10	US-09-764-853-661	Sequence 661, App
400	22	71.0	203	9	US-10-180-557-270	Sequence 270, App	473	22	71.0	291	10	US-09-870-501-1	Sequence 1, Appl1
401	22	71.0	203	9	US-10-173-700-270	Sequence 270, App	474	22	71.0	298	9	US-10-138-634-3	Sequence 3, Appl1
402	22	71.0	203	9	US-10-174-572-270	Sequence 270, App	475	22	71.0	298	10	US-09-925-301-1253	Sequence 1253, Ap
403	22	71.0	203	9	US-10-174-579-270	Sequence 270, App	476	22	71.0	300	10	US-09-815-242-11733	Sequence 11733, A
404	22	71.0	203	9	US-10-174-582-270	Sequence 270, App	477	22	71.0	300	10	US-09-815-242-13847	Sequence 13847, A
405	22	71.0	203	9	US-10-174-588-270	Sequence 270, App	478	22	71.0	301	8	US-08-902-572-14	Sequence 14, Appl
406	22	71.0	203	9	US-10-175-739-270	Sequence 270, App	479	22	71.0	301	10	US-09-863-179A-41	Sequence 41, Appl
407	22	71.0	203	9	US-10-175-740-270	Sequence 270, App	480	22	71.0	302	10	US-09-893-737-310	Sequence 310, App
408	22	71.0	203	9	US-10-175-743-270	Sequence 270, App	481	22	71.0	305	10	US-09-828-302-13	Sequence 13, Appl
409	22	71.0	203	9	US-10-176-488-270	Sequence 270, App	482	22	71.0	308	9	US-09-738-626-6317	Sequence 6317, Ap
410	22	71.0	203	9	US-10-176-92-270	Sequence 270, App	483	22	71.0	309	9	US-10-051-643-118	Sequence 118, App
411	22	71.0	203	9	US-10-176-747-270	Sequence 270, App	484	22	71.0	309	9	US-09-880-505-118	Sequence 118, App
412	22	71.0	203	9	US-10-176-750-270	Sequence 270, App	485	22	71.0	310	10	US-09-925-300-1602	Sequence 1602, Ap
413	22	71.0	203	9	US-10-176-985-270	Sequence 270, App	486	22	71.0	316	10	US-09-815-242-13294	Sequence 13294, A
414	22	71.0	203	9	US-10-176-987-270	Sequence 270, App	487	22	71.0	317	10	US-09-815-242-13524	Sequence 13524, A
415	22	71.0	203	9	US-10-176-991-270	Sequence 270, App	488	22	71.0	320	10	US-09-854-132-20	Sequence 20, Appl
416	22	71.0	203	9	US-10-176-992-270	Sequence 270, App	489	22	71.0	322	9	US-10-147-026-2	Sequence 2, Appl1
417	22	71.0	203	9	US-10-176-993-270	Sequence 270, App	490	22	71.0	323	10	US-09-815-242-5227	Sequence 5227, Ap
418	22	71.0	203	9	US-10-184-658-270	Sequence 270, App	491	22	71.0	324	10	US-09-815-242-12634	Sequence 12634, A
419	22	71.0	203	12	US-10-052-586-270	Sequence 270, App	492	22	71.0	327	10	US-09-815-242-13004	Sequence 13004, A
420	22	71.0	213	10	US-09-922-958-2	Sequence 2, Appl1	493	22	71.0	327	10	US-10-051-643-162	Sequence 162, App
421	22	71.0	213	10	US-09-755-016-2	Sequence 2, Appl1	494	22	71.0	327	9	US-09-880-505-162	Sequence 162, App
422	22	71.0	213	10	US-09-754-014-5	Sequence 5, Appl1	495	22	71.0	327	9	US-09-765-061B-75	Sequence 75, Appl
423	22	71.0	213	10	US-09-835-866-3	Sequence 3, Appl1	496	22	71.0	328	9	US-09-765-061B-76	Sequence 76, Appl
424	22	71.0	213	10	US-09-815-242-10461	Sequence 10461, A	497	22	71.0	328	10	US-09-777-404-29	Sequence 29, Appl
425	22	71.0	220	10	US-09-815-242-13988	Sequence 13988, A	498	22	71.0	328	10	US-09-777-404-29	Sequence 29, Appl
426	22	71.0	222	10	US-09-917-265-33	Sequence 33, Appl	499	22	71.0	329	10	US-09-804-969-4	Sequence 4, Appl1
427	22	71.0	222	10	US-09-917-265-47	Sequence 47, Appl	500	22	71.0	331	9	US-09-738-626-4960	Sequence 4960, Ap
428	22	71.0	222	10	US-09-917-265-105	Sequence 105, App	501	22	71.0	331	10	US-09-738-626-4864	Sequence 4864, App
429	22	71.0	222	12	US-10-079-616-3	Sequence 3, Appl1	502	22	71.0	333	10	US-09-851-194-4	Sequence 4, Appl1
430	22	71.0	226	9	US-09-976-673-8	Sequence 8, Appl1	503	22	71.0	336	10	US-09-939-484-2	Sequence 2, Appl1
431	22	71.0	226	9	US-09-976-673-10	Sequence 10, Appl1	504	22	71.0	336	10	US-09-939-484-2	Sequence 2, Appl1
432	22	71.0	226	9	US-09-976-673-26	Sequence 26, Appl1	505	22	71.0	339	9	US-09-981-876-246	Sequence 246, App
433	22	71.0	227	9	US-09-976-673-2	Sequence 2, Appl1	506	22	71.0	339	9	US-09-148-543-246	Sequence 246, App
434	22	71.0	227	9	US-09-976-673-4	Sequence 4, Appl1	507	22	71.0	342	9	US-09-712-363-210	Sequence 210, App
435	22	71.0	227	9	US-09-976-673-6	Sequence 6, Appl1	508	22	71.0	344	9	US-09-810-506-1	Sequence 1, Appl1
436	22	71.0	227	9	US-09-976-673-12	Sequence 12, Appl	509	22	71.0	346	10	US-09-925-301-1071	Sequence 1071, Ap
437	22	71.0	227	9	US-09-976-673-14	Sequence 14, Appl	510	22	71.0	348	10	US-09-755-016-4	Sequence 4, Appl1
438	22	71.0	227	9	US-09-976-673-24	Sequence 24, Appl	511	22	71.0	349	9	US-10-260-877-10	Sequence 10, Appl
439	22	71.0	223	12	US-10-038-107A-3	Sequence 3, Appl1	512	22	71.0	351	10	US-09-804-969-6	Sequence 6, Appl1
440	22	71.0	223	9	US-09-870-759-104	Sequence 104, App	513	22	71.0	357	10	US-09-823-414-9	Sequence 9, Appl1
441	22	71.0	238	10	US-09-823-356-1	Sequence 1, Appl1	514	22	71.0	359	10	US-09-888-615-114	Sequence 114, App
442	22	71.0	239	10	US-09-804-969-2	Sequence 2, Appl1	515	22	71.0	361	9	US-09-738-626-3735	Sequence 3735, Ap
443	22	71.0	239	10	US-09-804-969-10	Sequence 10, Appl	516	22	71.0	363	10	US-09-925-300-1530	Sequence 1530, Ap
444	22	71.0	241	9	US-10-074-956-29	Sequence 29, Appl	517	22	71.0	365	9	US-10-029-180-62	Sequence 62, Appl
445	22	71.0	247	9	US-09-738-626-3674	Sequence 3674, Ap	518	22	71.0	369	12	US-10-062-254-14	Sequence 14, Appl
446	22	71.0	248	10	US-09-853-450-28	Sequence 28, Appl	519	22	71.0	370	10	US-09-815-242-5328	Sequence 5328, Ap
447	22	71.0	248	10	US-09-815-242-12098	Sequence 12098, A	520	22	71.0	370	10	US-09-815-242-13553	Sequence 13553, A
448	22	71.0	250	9	US-09-853-450-30	Sequence 30, App	521	22	71.0	372	9	US-10-029-180-10	Sequence 10, Appl
449	22	71.0	250	9	US-09-738-626-9948	Sequence 9948, Ap	522	22	71.0	372	9	US-09-765-061B-74	Sequence 74, Appl
450	22	71.0	253	10	US-09-924-703-6	Sequence 6, Appl1	523	22	71.0	372	10	US-09-765-061B-76	Sequence 76, Appl
451	22	71.0	253	12	US-10-042-417-60	Sequence 60, Appl	524	22	71.0	372	10	US-09-766-984A-2	Sequence 2, Appl1
452	22	71.0	257	10	US-09-925-299-891	Sequence 891, App	525	22	71.0	374	10	US-09-815-242-13447	Sequence 13447, A
453	22	71.0	258	10	US-09-815-242-11725	Sequence 11725, A	526	22	71.0	374	10	US-09-815-242-13692	Sequence 13692, A
454	22	71.0	261	10	US-09-804-969-12	Sequence 12, Appl	527	22	71.0	375	10	US-09-815-242-10065	Sequence 10065, A
455	22	71.0	264	9	US-09-944-160-46	Sequence 46, Appl	528	22	71.0	377	10	US-09-815-242-13916	Sequence 13916, A
456	22	71.0	264	9	US-09-944-160-47	Sequence 47, Appl	529	22	71.0	383	10	US-09-864-762-48697	Sequence 48697, A
457	22	71.0	264	9	US-09-944-160-48	Sequence 48, Appl	530	22	71.0	384	9	US-09-924-400-333	Sequence 333, App

531	22	71.0	384	9	US-09-924-400-340	Sequence 340, App	604	22	71.0	517	10	US-09-909-903-21	Sequence 21, Appl
532	22	71.0	384	9	US-09-765-061B-72	Sequence 72, Appl	605	22	71.0	518	10	US-09-815-442-12279	Sequence 12279, A
533	22	71.0	384	9	US-09-765-061B-73	Sequence 73, Appl	606	22	71.0	522	10	US-09-765-772-120	Sequence 120, App
534	22	71.0	384	10	US-09-815-242-12644	Sequence 12644, A	607	22	71.0	523	9	US-10-042-141-53	Sequence 53, Appl
535	22	71.0	384	10	US-09-815-242-12883	Sequence 12883, A	608	22	71.0	523	9	US-10-051-663-114	Sequence 114, App
536	22	71.0	384	10	US-09-810-936-333	Sequence 333, App	609	22	71.0	523	9	US-09-880-505-114	Sequence 114, App
537	22	71.0	385	9	US-10-029-180-42	Sequence 42, Appl	610	22	71.0	523	10	US-09-726-643-53	Sequence 53, Appl
538	22	71.0	385	10	US-09-815-242-11483	Sequence 11483, A	611	22	71.0	526	10	US-09-828-825-2	Sequence 2, Appl
539	22	71.0	386	10	US-09-737-178-127	Sequence 127, App	612	22	71.0	527	10	US-09-828-825-2	Sequence 2, Appl
540	22	71.0	387	10	US-09-815-242-10416	Sequence 10416, A	613	22	71.0	528	10	US-09-815-442-5266	Sequence 5266, App
541	22	71.0	387	10	US-09-815-242-13750	Sequence 13750, A	614	22	71.0	533	10	US-09-917-265-39	Sequence 39, Appl
542	22	71.0	390	10	US-09-939-521-10	Sequence 10, Appl	615	22	71.0	533	10	US-09-917-265-62	Sequence 62, Appl
543	22	71.0	391	10	US-09-866-562-62	Sequence 62, Appl	616	22	71.0	537	10	US-09-738-666-4471	Sequence 4471, App
544	22	71.0	392	9	US-09-765-061B-77	Sequence 77, Appl	617	22	71.0	537	10	US-09-828-825-8	Sequence 8, Appl
545	22	71.0	392	10	US-09-815-242-5366	Sequence 5366, App	618	22	71.0	537	10	US-09-828-825-16	Sequence 16, Appl
546	22	71.0	393	10	US-09-784-077-2	Sequence 2, Appl	619	22	71.0	540	9	US-09-712-363-169	Sequence 169, App
547	22	71.0	398	10	US-09-815-242-13171	Sequence 13171, A	620	22	71.0	540	10	US-09-847-637B-6	Sequence 6, Appl
548	22	71.0	411	10	US-09-815-242-12606	Sequence 12606, A	621	22	71.0	541	9	US-10-051-663-160	Sequence 160, App
549	22	71.0	412	9	US-10-029-180-84	Sequence 84, Appl	622	22	71.0	541	9	US-09-880-505-160	Sequence 160, App
550	22	71.0	415	10	US-09-815-242-12099	Sequence 12099, A	623	22	71.0	542	10	US-09-828-825-6	Sequence 6, Appl
551	22	71.0	421	9	US-09-943-123-8	Sequence 8, Appl	624	22	71.0	542	10	US-09-828-825-14	Sequence 14, Appl
552	22	71.0	421	10	US-09-813-355-4	Sequence 4, Appl	625	22	71.0	543	10	US-09-927-738-2	Sequence 2, Appl
553	22	71.0	423	9	US-09-911-317-4	Sequence 4, Appl	626	22	71.0	547	10	US-09-828-825-10	Sequence 10, Appl
554	22	71.0	425	9	US-10-174-590-570	Sequence 570, App	627	22	71.0	547	10	US-09-828-825-12	Sequence 12, Appl
555	22	71.0	425	9	US-10-176-758-570	Sequence 570, App	628	22	71.0	549	9	US-09-959-845-2	Sequence 2, Appl
556	22	71.0	425	9	US-10-175-737-570	Sequence 570, App	629	22	71.0	552	10	US-09-835-854-2	Sequence 2, Appl
557	22	71.0	425	9	US-10-173-706-570	Sequence 570, App	630	22	71.0	555	9	US-10-024-809-2	Sequence 2, Appl
558	22	71.0	425	9	US-10-175-738-570	Sequence 570, App	631	22	71.0	574	9	US-09-990-385-13	Sequence 13, Appl
559	22	71.0	425	9	US-10-175-752-570	Sequence 570, App	632	22	71.0	577	10	US-09-815-442-10193	Sequence 10193, A
560	22	71.0	425	9	US-10-176-482-570	Sequence 570, App	633	22	71.0	577	10	US-09-815-442-11255	Sequence 11255, A
561	22	71.0	425	9	US-10-176-757-570	Sequence 570, App	634	22	71.0	577	10	US-09-815-442-13752	Sequence 13752, A
562	22	71.0	425	9	US-10-176-913-570	Sequence 570, App	635	22	71.0	580	9	US-09-928-457-37	Sequence 37, Appl
563	22	71.0	425	9	US-10-180-552-570	Sequence 570, App	636	22	71.0	582	10	US-09-916-658-4	Sequence 4, Appl
564	22	71.0	425	9	US-10-180-557-570	Sequence 570, App	637	22	71.0	582	10	US-08-801-196-27	Sequence 27, Appl
565	22	71.0	425	9	US-10-173-700-570	Sequence 570, App	638	22	71.0	582	10	US-09-919-97-84	Sequence 84, Appl
566	22	71.0	425	9	US-10-174-572-570	Sequence 570, App	639	22	71.0	583	10	US-09-815-442-10102	Sequence 10102, A
567	22	71.0	425	9	US-10-174-579-570	Sequence 570, App	640	22	71.0	583	10	US-09-815-442-14002	Sequence 14002, A
568	22	71.0	425	9	US-10-174-582-570	Sequence 570, App	641	22	71.0	587	10	US-09-815-442-12082	Sequence 12082, A
569	22	71.0	425	9	US-10-176-913-570	Sequence 570, App	642	22	71.0	590	10	US-09-893-817-2	Sequence 2, Appl
570	22	71.0	425	9	US-10-175-739-570	Sequence 570, App	643	22	71.0	599	10	US-09-952-113A-4	Sequence 4, Appl
571	22	71.0	425	9	US-10-175-743-570	Sequence 570, App	644	22	71.0	602	9	US-09-764-868-651	Sequence 651, App
572	22	71.0	425	9	US-10-176-488-570	Sequence 570, App	645	22	71.0	602	9	US-10-047-342-51	Sequence 51, Appl
573	22	71.0	425	9	US-10-176-492-570	Sequence 570, App	646	22	71.0	604	10	US-09-820-809-13	Sequence 13, Appl
574	22	71.0	425	9	US-10-176-747-570	Sequence 570, App	647	22	71.0	608	10	US-09-950-294-4	Sequence 4, Appl
575	22	71.0	425	9	US-10-176-750-570	Sequence 570, App	648	22	71.0	642	9	US-10-201-964-1	Sequence 1, Appl
576	22	71.0	425	9	US-10-176-985-570	Sequence 570, App	649	22	71.0	642	10	US-09-893-817-24	Sequence 24, Appl
577	22	71.0	425	9	US-10-176-987-570	Sequence 570, App	650	22	71.0	648	10	US-08-815-242-13995	Sequence 13995, A
578	22	71.0	425	9	US-10-176-987-570	Sequence 570, App	651	22	71.0	657	10	US-09-827-822-7	Sequence 7, Appl
579	22	71.0	425	9	US-10-176-991-570	Sequence 570, App	652	22	71.0	664	10	US-09-823-038A-48	Sequence 48, Appl
580	22	71.0	425	9	US-10-176-992-570	Sequence 570, App	653	22	71.0	669	10	US-09-801-196-28	Sequence 28, Appl
581	22	71.0	425	9	US-10-176-993-570	Sequence 570, App	654	22	71.0	690	9	US-10-068-059-10	Sequence 10, Appl
582	22	71.0	425	9	US-10-184-658-570	Sequence 570, App	655	22	71.0	709	9	US-10-068-059-8	Sequence 8, Appl
583	22	71.0	425	12	US-10-052-586-570	Sequence 570, App	656	22	71.0	709	10	US-09-815-242-5101	Sequence 5101, App
584	22	71.0	430	10	US-09-815-242-11315	Sequence 11315, A	657	22	71.0	720	9	US-10-047-412A-8	Sequence 8, Appl
585	22	71.0	432	10	US-09-815-242-10565	Sequence 10565, A	658	22	71.0	721	9	US-09-764-868B-731	Sequence 731, App
586	22	71.0	433	9	US-09-839-888-4	Sequence 4, Appl	659	22	71.0	724	9	US-10-068-059-12	Sequence 12, Appl
587	22	71.0	435	10	US-09-801-368-150	Sequence 150, App	660	22	71.0	731	9	US-09-738-626-4854	Sequence 4854, App
588	22	71.0	436	8	US-08-980-068B-19	Sequence 19, Appl	661	22	71.0	732	10	US-09-737-178-124	Sequence 124, App
589	22	71.0	437	9	US-09-943-123-14	Sequence 14, Appl	662	22	71.0	742	10	US-09-801-368B-26	Sequence 26, Appl
590	22	71.0	444	10	US-09-815-242-5482	Sequence 5482, App	663	22	71.0	746	9	US-10-068-059-6	Sequence 6, Appl
591	22	71.0	457	9	US-10-124-601-2	Sequence 2, Appl	664	22	71.0	746	9	US-09-982-107-4	Sequence 4, Appl
592	22	71.0	460	9	US-09-976-673-16	Sequence 16, Appl	665	22	71.0	756	10	US-09-771-621A-101	Sequence 101, App
593	22	71.0	470	9	US-09-976-673-18	Sequence 18, Appl	666	22	71.0	764	9	US-09-981-353-59	Sequence 59, Appl
594	22	71.0	471	9	US-09-738-626-3529	Sequence 3529, App	667	22	71.0	764	10	US-09-989-919-124	Sequence 124, App
595	22	71.0	473	10	US-09-815-242-12584	Sequence 12584, A	668	22	71.0	764	10	US-09-818-447-1	Sequence 1, Appl
596	22	71.0	478	9	US-09-859-888-2	Sequence 2, Appl	669	22	71.0	781	10	US-09-815-242-11269	Sequence 11269, A
597	22	71.0	501	10	US-09-815-242-11300	Sequence 11300, A	670	22	71.0	819	10	US-09-833-290-425	Sequence 425, App
598	22	71.0	502	10	US-09-323-998B-56	Sequence 56, Appl	671	22	71.0	831	9	US-09-738-447-1819	Sequence 1819, App
599	22	71.0	508	10	US-09-721-161A-102	Sequence 102, App	672	22	71.0	831	9	US-09-902-941-1819	Sequence 1819, App
600	22	71.0	511	10	US-09-917-265-44	Sequence 44, Appl	673	22	71.0	831	9	US-09-849-646-1819	Sequence 1819, App
601	22	71.0	511	10	US-09-917-265-67	Sequence 67, Appl	674	22	71.0	854	9	US-10-128-870-27	Sequence 27, Appl
602	22	71.0	515	9	US-09-738-626-4135	Sequence 4135, App	675	22	71.0	863	10	US-09-946-339-11	Sequence 11, Appl
603	22	71.0	517	10	US-09-796-138-21	Sequence 21, Appl	676	22	71.0	872	10	US-09-813-148-5	Sequence 5, Appl

677	22	71.0	885	10	US-09-815-242-5090	Sequence 5090, App	750	22	71.0	1177	10	US-09-873-873-28	Sequence 28, Appl
678	22	71.0	887	10	US-09-771-161A-192	Sequence 192, App	751	22	71.0	1177	10	US-09-873-873-34	Sequence 34, Appl
679	22	71.0	887	10	US-09-771-161A-193	Sequence 193, App	752	22	71.0	1178	10	US-09-851-194-2	Sequence 2, Appl
680	22	71.0	896	9	US-09-903-170C-5	Sequence 5, Appl	753	22	71.0	1193	10	US-09-873-873-30	Sequence 30, Appl
681	22	71.0	896	10	US-09-903-180B-5	Sequence 5, Appl	754	22	71.0	1221	10	US-09-738-626-513	Sequence 513, Ap
682	22	71.0	896	10	US-09-903-171A-5	Sequence 5, Appl	755	22	71.0	1221	10	US-09-919-891-2	Sequence 2, Appl
683	22	71.0	896	10	US-09-903-188A-5	Sequence 5, Appl	756	22	71.0	1229	10	US-09-815-244-5818	Sequence 5818, Ap
684	22	71.0	896	10	US-09-903-123A-5	Sequence 5, Appl	757	22	71.0	1229	10	US-09-815-244-1946	Sequence 12946, A
685	22	71.0	896	10	US-09-903-325A-5	Sequence 5, Appl	758	22	71.0	1241	10	US-09-759-623-2	Sequence 2, Appl
686	22	71.0	908	10	US-09-815-242-11847	Sequence 11847, A	759	22	71.0	1252	9	US-10-047-542-89	Sequence 89, Appl
687	22	71.0	912	10	US-09-817-664-2	Sequence 2, Appl	760	22	71.0	1274	9	US-10-020-215-2	Sequence 2, Appl
688	22	71.0	922	10	US-09-738-626-5476	Sequence 5476, App	761	22	71.0	1356	10	US-09-801-366-306	Sequence 306, App
689	22	71.0	922	10	US-09-886-468-19	Sequence 19, App	762	22	71.0	1352	9	US-09-738-626-4321	Sequence 4321, Ap
690	22	71.0	924	9	US-10-108-605-137	Sequence 137, App	763	22	71.0	1607	10	US-09-938-275-10	Sequence 10, Appl
691	22	71.0	941	9	US-09-993-598-353	Sequence 353, App	764	22	71.0	1609	10	US-09-938-275-11	Sequence 11, Appl
692	22	71.0	941	9	US-09-989-293A-353	Sequence 353, App	765	22	71.0	1633	9	US-09-738-626-6894	Sequence 6894, Ap
693	22	71.0	941	9	US-10-042-141-60	Sequence 60, Appl	766	22	71.0	1809	10	US-09-822-633-2	Sequence 2, Appl
694	22	71.0	941	9	US-09-989-735-353	Sequence 353, App	767	22	71.0	2249	9	US-09-865-557-2	Sequence 4, Appl
695	22	71.0	941	9	US-09-990-444-353	Sequence 353, App	768	22	71.0	26936	9	US-09-759-508-2	Sequence 2, Appl
696	22	71.0	941	9	US-09-989-730-353	Sequence 353, App	769	22	67.7	10	US-09-729-873-117	Sequence 117, App	
697	22	71.0	941	9	US-09-990-436-353	Sequence 353, App	770	22	67.7	13	US-09-774-639-277	Sequence 277, App	
698	22	71.0	941	9	US-09-991-181-353	Sequence 353, App	771	22	67.7	16	US-09-840-277-86	Sequence 82, Appl	
699	22	71.0	941	9	US-09-993-687-353	Sequence 353, App	772	22	67.7	16	US-09-840-277-86	Sequence 86, Appl	
700	22	71.0	941	9	US-09-989-734-353	Sequence 353, App	773	22	67.7	21	US-09-853-830-108	Sequence 108, App	
701	22	71.0	941	9	US-10-028-072-464	Sequence 464, App	774	22	67.7	32	US-09-861-764-28	Sequence 25, Appl	
702	22	71.0	941	9	US-09-993-667-353	Sequence 353, App	775	22	67.7	36	US-09-864-761-3869	Sequence 3869, A	
703	22	71.0	941	9	US-09-993-667-353	Sequence 353, App	776	22	67.7	36	US-09-733-524-9	Sequence 9, Appl	
704	22	71.0	941	9	US-10-121-049-464	Sequence 464, App	777	22	67.7	46	US-09-864-761-47826	Sequence 47826, A	
705	22	71.0	941	9	US-10-123-904-464	Sequence 464, App	778	22	67.7	47	US-09-864-761-49041	Sequence 49041, A	
706	22	71.0	941	9	US-10-140-470-464	Sequence 464, App	779	22	67.7	47	US-10-001-879-114	Sequence 114, App	
707	22	71.0	941	9	US-09-990-438-353	Sequence 353, App	780	22	67.7	53	US-09-843-676-20	Sequence 20, Appl	
708	22	71.0	941	9	US-09-990-562-353	Sequence 353, App	781	22	67.7	53	US-09-765-253-20	Sequence 20, Appl	
709	22	71.0	941	9	US-09-997-428-353	Sequence 353, App	782	22	67.7	53	US-09-438-486-20	Sequence 20, Appl	
710	22	71.0	941	9	US-09-997-666-353	Sequence 353, App	783	22	67.7	58	US-09-864-761-40567	Sequence 40567, A	
711	22	71.0	941	9	US-10-175-746-464	Sequence 464, App	784	22	67.7	61	US-08-808-031A-19	Sequence 19, Appl	
712	22	71.0	941	9	US-10-176-918-464	Sequence 464, App	785	22	67.7	63	US-09-864-761-37193	Sequence 37193, A	
713	22	71.0	941	10	US-10-176-921-464	Sequence 464, App	786	22	67.7	68	US-09-895-913A-272	Sequence 272, App	
714	22	71.0	941	10	US-09-726-643-60	Sequence 60, Appl	787	22	67.7	70	US-09-974-879-191	Sequence 191, App	
715	22	71.0	941	10	US-09-989-723-353	Sequence 353, App	788	22	67.7	70	US-09-864-761-34690	Sequence 34690, A	
716	22	71.0	941	10	US-09-989-723-353	Sequence 353, App	789	22	67.7	71	US-09-864-761-36979	Sequence 36979, A	
717	22	71.0	941	10	US-09-989-279-353	Sequence 353, App	790	22	67.7	72	US-09-925-301-1505	Sequence 1505, Ap	
718	22	71.0	941	10	US-09-989-727-353	Sequence 353, App	791	22	67.7	72	US-09-808-212A-8	Sequence 8, Appl	
719	22	71.0	941	10	US-09-989-731-353	Sequence 353, App	792	22	67.7	72	US-09-808-212A-10	Sequence 10, Appl	
720	22	71.0	941	10	US-09-989-732-353	Sequence 353, App	793	22	67.7	73	US-09-939-980-416	Sequence 416, App	
721	22	71.0	941	10	US-09-991-073-353	Sequence 353, App	794	22	67.7	75	US-09-808-212A-18	Sequence 18, Appl	
722	22	71.0	941	10	US-09-990-442-353	Sequence 353, App	795	22	67.7	77	US-09-939-980-511	Sequence 511, App	
723	22	71.0	941	10	US-09-991-163-353	Sequence 353, App	796	22	67.7	82	US-09-808-212A-2	Sequence 2, Appl	
724	22	71.0	941	10	US-09-993-604-353	Sequence 353, App	797	22	67.7	83	US-10-042-141-84	Sequence 84, Appl	
725	22	71.0	941	10	US-09-990-556-353	Sequence 353, App	798	22	67.7	83	US-09-726-643-84	Sequence 84, Appl	
726	22	71.0	941	10	US-09-989-721-353	Sequence 353, App	799	22	67.7	84	US-09-974-879-289	Sequence 289, App	
727	22	71.0	961	10	US-09-220-091-3	Sequence 3, Appl	800	22	67.7	85	US-09-730-611-35	Sequence 35, Appl	
728	22	71.0	978	9	US-09-938-901-8	Sequence 8, Appl	801	22	67.7	87	US-09-861-550-36	Sequence 36, Appl	
729	22	71.0	979	10	US-09-903-187A-5	Sequence 5, Appl	802	22	67.7	91	US-09-764-868-691	Sequence 691, App	
730	22	71.0	990	12	US-10-047-676A-7	Sequence 7, Appl	803	22	67.7	95	US-09-764-868-784	Sequence 784, App	
731	22	71.0	993	10	US-09-815-242-5809	Sequence 5809, App	804	22	67.7	95	US-09-764-868-1189	Sequence 1189, Ap	
732	22	71.0	993	10	US-09-881-752A-80	Sequence 80, Appl	805	22	67.7	105	US-09-738-626-3621	Sequence 3621, Ap	
733	22	71.0	999	9	US-09-895-913A-226	Sequence 226, App	806	22	67.7	109	US-09-925-299-1043	Sequence 1043, Ap	
734	22	71.0	1002	10	US-09-815-242-12899	Sequence 12899, A	807	22	67.7	109	US-09-220-200A-7	Sequence 7, Appl	
735	22	71.0	1002	10	US-09-815-242-13158	Sequence 13158, A	808	22	67.7	115	US-09-738-626-406A-7	Sequence 406A-7	
736	22	71.0	1040	10	US-09-765-772-118	Sequence 118, App	809	22	67.7	115	US-09-265-959-352	Sequence 352, App	
737	22	71.0	1042	10	US-09-815-242-13345	Sequence 13345, A	810	22	67.7	117	US-09-864-761-37175	Sequence 37175, A	
738	22	71.0	1042	10	US-09-825-697-2	Sequence 2, Appl	811	22	67.7	119	US-08-945-459A-1	Sequence 1, Appl	
739	22	71.0	1135	10	US-09-756-643-2	Sequence 2, Appl	812	22	67.7	119	US-09-880-708-13	Sequence 13, Appl	
740	22	71.0	1156	10	US-09-826-660-15	Sequence 15, Appl	813	22	67.7	119	US-09-068-253-2	Sequence 2, Appl	
741	22	71.0	1160	10	US-09-815-242-5480	Sequence 5480, Ap	814	22	67.7	119	US-09-864-761-43653	Sequence 43653, A	
742	22	71.0	1168	10	US-09-815-242-12151	Sequence 12151, A	815	22	67.7	120	US-09-945-182A-4	Sequence 4, Appl	
743	22	71.0	1169	10	US-09-815-242-13448	Sequence 13448, A	816	22	67.7	122	US-10-115-406-7	Sequence 7, Appl	
744	22	71.0	1170	10	US-09-945-265-2	Sequence 2, Appl	817	22	67.7	123	US-09-859-211-36	Sequence 36, Appl	
745	22	71.0	1170	10	US-09-873-873-10	Sequence 10, Appl	818	22	67.7	123	US-09-815-459-7	Sequence 7, Appl	
746	22	71.0	1177	10	US-09-873-873-12	Sequence 12, Appl	819	22	67.7	131	US-09-738-626-5916	Sequence 5916, Ap	
747	22	71.0	1177	10	US-09-873-873-14	Sequence 14, Appl	820	22	67.7	135	US-09-764-868-1098	Sequence 1098, Ap	
748	22	71.0	1177	10	US-09-873-873-14	Sequence 14, Appl	821	22	67.7	140	US-09-730-617-46	Sequence 46, Appl	
749	22	71.0	1177	10	US-09-873-873-26	Sequence 26, Appl	822	22	67.7	142	US-09-797-481-2	Sequence 2, Appl	

823	21	67.7	142	10	US-09-844-736-4	Sequence 4, Appl1	896	21	67.7	312	10	US-09-815-242-5267	Sequence 5367, Ap
824	21	67.7	144	9	US-09-929-230-8	Sequence 8, Appl1	897	21	67.7	316	10	US-09-462-846-5	Sequence 5, Appl1
825	21	67.7	144	10	US-09-730-617-45	Sequence 45, Appl	898	21	67.7	321	9	US-09-945-182-26	Sequence 26, Appl
826	21	67.7	154	9	US-10-090-185-11	Sequence 24, Appl	899	21	67.7	321	10	US-09-886-055-259	Sequence 259, Ap
827	21	67.7	155	10	US-09-425-021-24	Sequence 11, Appl	900	21	67.7	323	10	US-09-886-055-77	Sequence 77, Appl
828	21	67.7	174	9	US-09-764-868-1131	Sequence 613, Ap	901	21	67.7	323	10	US-09-771-730-8	Sequence 8, Appl1
829	21	67.7	175	10	US-09-925-297-613	Sequence 297, App	902	21	67.7	323	10	US-09-771-730-16	Sequence 16, Appl
830	21	67.7	177	10	US-09-741-669-293	Sequence 317, App	903	21	67.7	323	10	US-09-771-730-18	Sequence 18, Appl
831	21	67.7	177	10	US-09-912-020-317	Sequence 317, App	904	21	67.7	323	10	US-09-771-730-18	Sequence 18, Appl
832	21	67.7	177	10	US-09-815-242-5164	Sequence 5164, Ap	905	21	67.7	328	9	US-09-735-055-31	Sequence 31, Appl
833	21	67.7	177	10	US-09-815-242-11112	Sequence 10344, A	906	21	67.7	342	10	US-09-815-242-13188	Sequence 13188, A
834	21	67.7	177	10	US-09-815-242-11112	Sequence 11112, A	907	21	67.7	343	10	US-09-849-9808-1	Sequence 1, Appl1
835	21	67.7	179	9	US-10-223-047-11	Sequence 14103, A	908	21	67.7	346	10	US-09-864-761-16084	Sequence 36084, A
836	21	67.7	179	9	US-09-877-964-4	Sequence 11, Appl	909	21	67.7	350	10	US-09-791-961-5	Sequence 5, Appl1
837	21	67.7	180	10	US-09-765-272-186	Sequence 186, App	910	21	67.7	352	10	US-09-815-242-11287	Sequence 1, Appl1
838	21	67.7	182	10	US-09-815-242-12316	Sequence 12316, A	911	21	67.7	353	10	US-09-815-242-11287	Sequence 11287, A
839	21	67.7	182	10	US-09-975-374A-10	Sequence 10, Appl	912	21	67.7	354	9	US-10-121-857-38	Sequence 38, Appl
840	21	67.7	183	10	US-09-815-242-5833	Sequence 5833, Ap	913	21	67.7	357	10	US-09-740-046-8	Sequence 38, Appl
841	21	67.7	190	10	US-09-877-964-2	Sequence 2, Appl1	914	21	67.7	358	10	US-09-815-242-12380	Sequence 8, Appl1
842	21	67.7	190	10	US-09-877-964-2	Sequence 4, Appl1	915	21	67.7	358	10	US-09-925-637-4	Sequence 12380, A
843	21	67.7	190	10	US-09-877-964-4	Sequence 2, Appl1	916	21	67.7	359	12	US-10-072-152-2	Sequence 4, Appl1
844	21	67.7	197	9	US-09-905-558C-2	Sequence 32, Appl	917	21	67.7	360	9	US-09-738-626-4558	Sequence 2, Appl1
845	21	67.7	198	9	US-10-140-293-32	Sequence 1216, Ap	918	21	67.7	360	10	US-09-815-242-11186	Sequence 4958, Ap
846	21	67.7	199	9	US-09-796-692-1216	Sequence 1675, Ap	919	21	67.7	361	9	US-09-815-242-11186	Sequence 11186, Ap
847	21	67.7	199	9	US-09-796-692-1675	Sequence 2233, Ap	920	21	67.7	363	10	US-09-815-242-12579	Sequence 4650, Ap
848	21	67.7	199	9	US-10-140-293-31	Sequence 31, Appl	921	21	67.7	365	10	US-09-815-242-11917	Sequence 12579, A
849	21	67.7	204	10	US-09-815-242-13428	Sequence 13428, A	922	21	67.7	371	10	US-09-925-637-4	Sequence 1681, Ap
850	21	67.7	204	10	US-09-815-242-13687	Sequence 13687, A	923	21	67.7	373	10	US-09-813-398-16	Sequence 36, Appl
851	21	67.7	206	10	US-09-935-428A-10	Sequence 10, Appl	924	21	67.7	373	10	US-09-905-173-26	Sequence 26, Appl
852	21	67.7	210	10	US-09-997-165-6	Sequence 6, Appl	925	21	67.7	380	10	US-09-815-242-12770	Sequence 26, Appl
853	21	67.7	220	9	US-09-860-670-129	Sequence 129, App	926	21	67.7	381	10	US-09-738-626-6596	Sequence 12770, A
854	21	67.7	220	9	US-09-764-868-949	Sequence 949, App	927	21	67.7	382	10	US-09-815-242-5342	Sequence 6596, Ap
855	21	67.7	220	9	US-09-764-868-1331	Sequence 1231, Ap	928	21	67.7	382	10	US-09-815-242-13912	Sequence 5342, Ap
856	21	67.7	220	9	US-09-997-165-2	Sequence 5, Appl1	929	21	67.7	384	9	US-09-924-400-326	Sequence 334, Ap
857	21	67.7	220	10	US-09-770-940-5	Sequence 78, Appl	930	21	67.7	384	9	US-09-924-400-326	Sequence 334, Ap
858	21	67.7	221	10	US-09-765-272-78	Sequence 9, Appl1	931	21	67.7	384	9	US-09-924-400-332	Sequence 332, App
859	21	67.7	221	10	US-10-170-102-9	Sequence 11, Appl	932	21	67.7	384	10	US-09-924-400-332	Sequence 332, App
860	21	67.7	224	10	US-09-765-213A-11	Sequence 30, Appl	933	21	67.7	384	10	US-09-825-301-8	Sequence 8, Appl1
861	21	67.7	240	10	US-09-945-182-30	Sequence 2, Appl1	934	21	67.7	384	10	US-09-825-301-8	Sequence 11956, A
862	21	67.7	240	10	US-09-997-165-2	Sequence 1694, A	935	21	67.7	384	10	US-09-815-242-11196	Sequence 5188, Ap
863	21	67.7	242	10	US-09-815-242-11694	Sequence 3952, Ap	936	21	67.7	384	10	US-09-815-242-11196	Sequence 304, App
864	21	67.7	248	10	US-09-725-178-2	Sequence 2, Appl1	937	21	67.7	384	10	US-09-810-936-326	Sequence 326, App
865	21	67.7	248	10	US-09-839-185-10	Sequence 10, Appl	938	21	67.7	384	10	US-09-810-936-332	Sequence 332, App
866	21	67.7	258	9	US-09-870-759-14	Sequence 14, Appl	939	21	67.7	384	10	US-09-810-936-332	Sequence 334, App
867	21	67.7	261	10	US-09-815-242-3376	Sequence 5376, Ap	940	21	67.7	384	10	US-09-429-553-304	Sequence 304, App
868	21	67.7	263	9	US-09-945-182-32	Sequence 32, Appl	941	21	67.7	388	9	US-09-945-182-34	Sequence 34, Appl
869	21	67.7	263	10	US-09-859-882-4	Sequence 4, Appl1	942	21	67.7	390	10	US-09-815-242-12402	Sequence 12402, A
870	21	67.7	264	10	US-09-726-899-1	Sequence 1, Appl1	943	21	67.7	392	9	US-09-738-626-1635	Sequence 336, App
871	21	67.7	266	10	US-09-726-899-9	Sequence 9, Appl1	944	21	67.7	394	9	US-09-924-400-336	Sequence 12007, A
872	21	67.7	266	10	US-09-815-242-10197	Sequence 10197, A	945	21	67.7	394	10	US-09-815-242-12007	Sequence 28, Appl
873	21	67.7	266	10	US-09-815-242-12291	Sequence 12291, A	946	21	67.7	411	9	US-09-945-182-38	Sequence 13330, A
874	21	67.7	268	10	US-09-925-637-60	Sequence 60, Appl	947	21	67.7	419	10	US-09-815-242-13330	Sequence 28, Appl
875	21	67.7	276	9	US-09-764-868-811	Sequence 811, App	948	21	67.7	422	10	US-09-905-119-2	Sequence 2, Appl1
876	21	67.7	276	9	US-09-738-626-5519	Sequence 5519, Ap	949	21	67.7	423	10	US-09-798-029-8	Sequence 8, Appl1
877	21	67.7	282	10	US-09-815-242-14057	Sequence 14057, A	950	21	67.7	425	10	US-09-828-313-29	Sequence 29, Appl
878	21	67.7	289	10	US-09-10-16-253-7	Sequence 7, Appl1	951	21	67.7	425	10	US-09-828-313-29	Sequence 10, Appl
879	21	67.7	290	10	US-09-815-242-11940	Sequence 11940, A	952	21	67.7	434	1	US-08-325-783-610	Sequence 3, Appl1
880	21	67.7	290	10	US-09-815-242-5743	Sequence 5743, Ap	953	21	67.7	434	9	US-09-738-626-610	Sequence 4610, Ap
881	21	67.7	293	9	US-09-815-242-12742	Sequence 12742, A	954	21	67.7	434	10	US-09-731-8266-4	Sequence 4610, Ap
882	21	67.7	294	9	US-09-945-182-2	Sequence 2, Appl1	955	21	67.7	435	10	US-09-815-242-5702	Sequence 5702, Ap
883	21	67.7	296	10	US-09-960-472-1	Sequence 1, Appl1	956	21	67.7	441	10	US-09-925-301-1303	Sequence 1303, Ap
884	21	67.7	299	10	US-09-771-730-85	Sequence 85, Appl	957	21	67.7	443	10	US-09-929-060-1	Sequence 1, Appl1
885	21	67.7	299	10	US-09-771-730-90	Sequence 90, Appl	958	21	67.7	443	10	US-09-929-060-2	Sequence 2, Appl1
886	21	67.7	302	10	US-09-854-122-19	Sequence 19, Appl	959	21	67.7	446	10	US-09-747-116-2	Sequence 12, Appl
887	21	67.7	302	10	US-09-854-122-19	Sequence 19, Appl	960	21	67.7	446	10	US-09-910-186A-12	Sequence 339, App
888	21	67.7	303	10	US-09-771-730-61	Sequence 61, Appl	961	21	67.7	451	9	US-09-866-356-4	Sequence 4, Appl1
889	21	67.7	303	10	US-09-771-730-92	Sequence 92, Appl	962	21	67.7	452	10	US-09-764-870-339	Sequence 339, App
890	21	67.7	305	10	US-08-325-278-1	Sequence 1, Appl1	963	21	67.7	454	9	US-09-738-626-5211	Sequence 5211, Ap
891	21	67.7	305	10	US-08-325-278-1	Sequence 2, Appl1	964	21	67.7	459	10	US-09-866-987-8	Sequence 8, Appl1
892	21	67.7	305	10	US-08-325-278-1	Sequence 266, Appl	965	21	67.7	460	10	US-09-841-880-2	Sequence 35, Appl
893	21	67.7	310	9	US-09-801-568-286	Sequence 286, Ap	966	21	67.7	461	10	US-09-862-027-35	Sequence 35, Appl
894	21	67.7	312	9	US-09-738-626-3669	Sequence 6812, Ap	967	21	67.7	461	10	US-09-862-027-35	Sequence 35, Appl
895	21	67.7	312	10	US-09-788-626-21	Sequence 21, Appl	968	21	67.7	463	9	US-09-738-626-3970	Sequence 3970, Ap

```
569 21 67.7 467 10 US-09-929-060-3 Sequence 3, Appli
970 21 67.7 467 12 US-10-083-452-11 Sequence 11, Appl
971 21 67.7 477 9 US-09-738-626-6985 Sequence 6985, Ap
972 21 67.7 479 10 US-09-881-752A-14 Sequence 14, Appl
973 21 67.7 481 10 US-09-817-464-8 Sequence 8, Appl
974 21 67.7 485 10 US-09-815-242-5503 Sequence 5503, Ap
975 21 67.7 491 9 US-09-895-913A-106 Sequence 106, App
976 21 67.7 491 10 US-09-815-242-12492 Sequence 12492, A
977 21 67.7 491 10 US-09-815-242-12864 Sequence 12864, A
978 21 67.7 494 10 US-09-815-242-11055 Sequence 11055, A
979 21 67.7 495 9 US-09-880-708-10 Sequence 10, Appl
980 21 67.7 501 6 US-09-784-249-2 Sequence 2, Appl
981 21 67.7 501 6 US-09-981-490B-1 Sequence 1, Appl
982 21 67.7 501 10 US-09-730-772-13 Sequence 13, Appl
983 21 67.7 501 10 US-09-735-849-13 Sequence 13, Appl
984 21 67.7 502 10 US-09-813-398-37 Sequence 37, Appl
985 21 67.7 502 10 US-09-815-242-11205 Sequence 11205, A
986 21 67.7 504 10 US-09-815-242-12537 Sequence 12537, A
987 21 67.7 509 9 US-09-258-031B-75 Sequence 75, Appl
988 21 67.7 511 9 US-10-121-032-21 Sequence 21, Appl
989 21 67.7 511 9 US-09-941-947A-22 Sequence 22, Appl
990 21 67.7 511 10 US-09-934-903-16 Sequence 16, Appl
991 21 67.7 511 10 US-09-934-868-76 Sequence 76, Appl
992 21 67.7 526 9 US-09-738-626-6454 Sequence 6454, Ap
993 21 67.7 528 9 US-10-121-498-12 Sequence 12, Appl
994 21 67.7 529 9 US-09-924-400-324 Sequence 324, App
995 21 67.7 529 9 US-08-258-031B-16 Sequence 16, Appl
996 21 67.7 529 9 US-08-258-031B-20 Sequence 20, Appl
997 21 67.7 529 10 US-09-810-936-324 Sequence 324, App
998 21 67.7 533 9 US-09-905-291A-332 Sequence 332, App
999 21 67.7 533 9 US-09-902-853-332 Sequence 332, App
1000 21 67.7 533 9 US-09-907-824-332 Sequence 332, App
```

ALIGNMENTS

```
RESULT 1
US-09-816-825-9
; Sequence 9, Application US/09816825
; Patient No. US20010051370A1
; GENERAL INFORMATION:
; APPLICANT: Bistrup, Annette
; APPLICANT: Rosen, Steven D.
; APPLICANT: Hemmerich, Stefan
; TITLE OF INVENTION: GLYCOSYL SULFORANSFERASE-3
; FILE REFERENCE: 6510-107CON
; CURRENT APPLICATION NUMBER: US/09/816,825
; CURRENT FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/045,284
; PRIOR FILING DATE: 1998-03-20
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-816-825-9

Query Match          100.0%; Score 31; DB 10; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.2e+05;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-9
```

```
Query Match          100.0%; Score 31; DB 10; Length 169;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
DB 127 VRYEDL 132
```

```
RESULT 3
US-09-927-602-8
; Sequence 8, Application US/09927602
; Patient No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 171
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-8
```

```
Query Match          100.0%; Score 31; DB 10; Length 171;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 VRYEDL 6
DB 129 VRYEDL 134
```

```
RESULT 4
US-09-927-602-10
; Sequence 10, Application US/09927602
; Patient No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 10
LENGTH: 179
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-10

Query Match 100.0%; Score 31; DB 10; Length 179;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 137 VRYEDL 142

RESULT 5
US-10-007-262-1
Sequence 1, Application US/10007262
Patent No. US20020164748A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Tangemann, Kirsten
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107C1P
CURRENT APPLICATION NUMBER: US/10/007,262
CURRENT FILING DATE: 2001-11-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/190,911
PRIOR FILING DATE: EARLIER FILING DATE: 1998-11-12
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 1
LENGTH: 386
TYPE: PRT
ORGANISM: H. sapiens
US-10-007-262-1

Query Match 100.0%; Score 31; DB 9; Length 386;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 274 VRYEDL 279

RESULT 6
US-09-816-825-2
Sequence 2, Application US/09816825
Patent No. US20010051370A1
GENERAL INFORMATION:
APPLICANT: Bistrup, Annette
APPLICANT: Rosen, Steven D.
APPLICANT: Hemmerich, Stefan
TITLE OF INVENTION: GLYCOSYL SULFOTRANSFERASE-3
FILE REFERENCE: 6510-107CON
CURRENT APPLICATION NUMBER: US/09/816,825
CURRENT FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/045,284
PRIOR FILING DATE: 1998-03-20
NUMBER OF SEQ ID NOS: 9
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 386
TYPE: PRT
ORGANISM: Homo sapiens
US-09-816-825-2

Query Match 100.0%; Score 31; DB 10; Length 386;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 274 VRYEDL 279

RESULT 7
US-09-927-602-5
Sequence 5, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
Dystrophy
FILE REFERENCE: P-LT 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 5
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-927-602-5

Query Match 100.0%; Score 31; DB 10; Length 418;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 294 VRYEDL 299

RESULT 8
US-10-212-933-2
Sequence 2, Application US/10212933
Publication No. US20030008366A1
GENERAL INFORMATION:
APPLICANT: Uchimura, Kenji
APPLICANT: Muramatsu, Hideki
APPLICANT: Kadomatsu, Kenji
APPLICANT: Kannaqi, Reiji
APPLICANT: Hadochi, Osami
APPLICANT: Muramatsu, Takashi
TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
TITLE OF INVENTION: SULFOTRANSFERASE AND
FILE REFERENCE: DNA ENCODING THE SAME
FILE REFERENCE: TOYAMA1.001AUS
CURRENT APPLICATION NUMBER: US/10/212,933
CURRENT FILING DATE: 2002-08-05
PRIOR APPLICATION NUMBER: US/09/263,023
PRIOR FILING DATE: 1999-03-05
PRIOR APPLICATION NUMBER: JP 10-54007
PRIOR FILING DATE: 1998-03-05
PRIOR APPLICATION NUMBER: JP 10-177844
PRIOR FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 10
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 483
TYPE: PRT
ORGANISM: Mus musculus
US-10-212-933-2

Query Match 100.0%; Score 31; DB 9; Length 483;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 376 VREEDL 381

RESULT 9
US-10-212-933-4
; Sequence 4, Application US/10212933
; Publication No. US20030008366A1
; GENERAL INFORMATION:
; APPLICANT: Uchimura, Kenji
; APPLICANT: Muramatsu, Hideki
; APPLICANT: Kadamatsu, Kenji
; APPLICANT: Kannagi, Reiji
; APPLICANT: Habuchi, Osami
; APPLICANT: Muramatsu, Takashi
; TITLE OF INVENTION: POLYPEPTIDE OF N-ACETYLGLUCOSAMINE-6-0-
; TITLE OF INVENTION: SULFOTRANSFERASE AND
; FILE REFERENCE: TOYAMA41.001AUS
; CURRENT APPLICATION NUMBER: US/10/212,933
; PRIOR FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: US/09/263,023
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: JP 10-54007
; PRIOR FILING DATE: 1998-03-05
; PRIOR APPLICATION NUMBER: JP 10-177844
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-933-4

Query Match 100.0%; Score 31; DB 9; Length 484;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
Db 377 VREEDL 382

RESULT 10
US-09-833-790-255
; Sequence 255, Application US/09833790
; Patent No. US20020068288A1
; GENERAL INFORMATION:
; APPLICANT: Lodges, Michael J.
; APPLICANT: Wang, Tonglong
; APPLICANT: Secrist, Heather
; APPLICANT: Mohamath, Raedch
; APPLICANT: Indrias, Carol Y.
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.512
; CURRENT APPLICATION NUMBER: US/09/833,790
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 255
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-833-790-255

Query Match 100.0%; Score 31; DB 10; Length 531;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||

Db 424 VREEDL 429

RESULT 11
US-09-927-602-11
; Sequence 11, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; TITLE OF INVENTION: Dystrophy
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 174
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-11

Query Match 90.3%; Score 28; DB 10; Length 174;
Best Local Similarity 83.3%; Pred. No. 44;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
Db 132 VREEDV 137

RESULT 12
US-10-000-864-6
; Sequence 6, Application US/10000864
; Patent No. US20020146798A1
; GENERAL INFORMATION:
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION
; TITLE OF INVENTION: HUMAN MEK1 PROTEIN AND NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND USES THEREFOR
; FILE REFERENCE: CRI-085CPC
; CURRENT APPLICATION NUMBER: US/10/000,864
; CURRENT FILING DATE: 2001-10-31
; EARLIER APPLICATION NUMBER: 09/423,890
; EARLIER FILING DATE: 2000-06-03
; EARLIER APPLICATION NUMBER: PCT/US99/05556
; EARLIER FILING DATE: 1999-03-15
; EARLIER APPLICATION NUMBER: USSN 60/078,153
; EARLIER FILING DATE: 1998-03-16
; EARLIER APPLICATION NUMBER: USSN 60/099,165
; EARLIER FILING DATE: 1998-09-04
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 626
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-000-864-6

Query Match 90.3%; Score 28; DB 12; Length 626;
Best Local Similarity 83.3%; Pred. No. 1,6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VREEDL 6
|||||
Db 64 VREEDV 69

RESULT 13
US-10-000-864-12
; Sequence 12, Application US/10000864

Patent No. US20020146798A1
GENERAL INFORMATION:
APPLICANT: CADUS PHARMACEUTICAL CORPORATION
TITLE OF INVENTION: HUMAN MEK1 PROTEIN AND NUCLEIC ACID MOLECULES
TITLE OF INVENTION: AND USES THEREFOR
FILE REFERENCE: CPT-0850PC
CURRENT APPLICATION NUMBER: US/10/000,864
EARLIER FILING DATE: 2001-10-31
EARLIER APPLICATION NUMBER: 09/423,890
EARLIER FILING DATE: 2000-06-03
EARLIER APPLICATION NUMBER: PCT/US99/05556
EARLIER FILING DATE: 1999-03-15
EARLIER APPLICATION NUMBER: USSN 60/078,153
EARLIER FILING DATE: 1998-03-16
EARLIER APPLICATION NUMBER: USSN 60/099,165
EARLIER FILING DATE: 1998-09-04
NUMBER OF SEQ ID NOS: 38
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 626
TYPE: PRT
ORGANISM: Mus musculus
US-10-000-864-12

Query Match
Best Local Similarity 90.3%; Score 28; DB 12; Length 626;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 1.6e+02;

QY 1 VRYEDL 6
|||||
DB 64 VRYEDV 69

RESULT 14
US-09-864-761-35570
Sequence 35570, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annotmax Sequence Listing Engine vers. 1.1
SEQ ID NO 35570
LENGTH: 69
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096700.6
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.94
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.8
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.9
OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.8
OTHER INFORMATION: EST_HUMAN HIT: D4816.1, EVALU8 5.00e-35
OTHER INFORMATION: SWISSPROT HIT: P40421, EVALU8 7.00e-04
US-09-864-761-35570

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 69;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Pred. No. 28;

QY 2 RYEDL 6
|||||
DB 56 RYEDL 60

RESULT 15
US-09-864-761-49114
Sequence 49114, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
FILE REFERENCE: Aeomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Autonomax Sequence Listing Engine vers. 1.1
SEQ ID NO 49114
LENGTH: 69
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AL096700.14
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.8
OTHER INFORMATION: SWISSPROT HIT: P40421, EVALUATE 7.00e-04
OTHER INFORMATION: EST_HUMAN HIT: D44816.1, EVALUATE 8.00e-35
US-09-864-761-49114

Query Match 87.1%; Score 27; DB 10; Length 69;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
DB 56 RYEDL 60

RESULT 16
US-09-738-626-3600
Sequence 3600, Application US/09738626
Publication No. US20020197605A1
GENERAL INFORMATION:
APPLICANT: NAKAGAWA, SATOSHI
APPLICANT: MIZOGUCHI, HIROSHI
APPLICANT: ANDO, SEIRO
APPLICANT: HAYASHI, MIKIRO
APPLICANT: OCHIAI, KEIKO
APPLICANT: YOKOI, HARUHIKO
APPLICANT: TATEISHI, NAOKO
APPLICANT: SENOH, AKIHIRO
APPLICANT: IKEDA, MASATO
APPLICANT: OZAKI, AKIO
TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
FILE REFERENCE: 249-125
CURRENT APPLICATION NUMBER: US/09/738,626
CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: PatentIn ver. 3.0
SEQ ID NO 3600
LENGTH: 157
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-3600

Query Match 87.1%; Score 27; DB 9; Length 157;
Best Local Similarity 83.3%; Pred. No. 65;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 25 VRYEDL 30

RESULT 17
US-09-927-602-6
Sequence 6, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-6

Query Match 87.1%; Score 27; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 127 VRYEDL 132

RESULT 18
US-09-927-602-7
Sequence 7, Application US/09927602
Patent No. US20020061562A1
GENERAL INFORMATION:
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
TITLE OF INVENTION: Methods of Treating Macular Corneal
FILE REFERENCE: P-LJ 4852
CURRENT APPLICATION NUMBER: US/09/927,602
CURRENT FILING DATE: 2001-08-09
PRIOR APPLICATION NUMBER: US 09/638,211
PRIOR FILING DATE: 2000-08-11
NUMBER OF SEQ ID NOS: 38
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7
LENGTH: 169
TYPE: PRT
ORGANISM: Homo Sapien
US-09-927-602-7

Query Match 87.1%; Score 27; DB 10; Length 169;
Best Local Similarity 83.3%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
DB 127 VRYEDL 132

RESULT 19
US-09-361-741-2

```
; Sequence 2, Application US/09361741
; Patient No. US20020048784A1
; GENERAL INFORMATION:
; APPLICANT: HUDSON, PETER L
; APPLICANT: ROSEN, CRAIG A
; APPLICANT: HE, WEI WU
; TITLE OF INVENTION: PROSTATIC GROWTH FACTOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN, CECCHI,
; ADDRESSEE: STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NJ
; COUNTRY: US
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,741
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,607
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: FERRARO, GREGORY D
; REGISTRATION NUMBER: 36,134
; REFERENCE/DOCKET NUMBER: 325800-329
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201)994-1700
; TELEFAX: (201)994-1744
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 295 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-361-741-2

Query Match      87.1%; Score 27; DB 10; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      46 RYEDL 50

RESULT 20
; Sequence 188, Application US/09981353
; Patient No. US20020160382A1
; GENERAL INFORMATION:
; APPLICANT: Lasek, Amy W,
; APPLICANT: Jones, David A.
; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
; FILE REFERENCE: PA-0038 US
; CURRENT APPLICATION NUMBER: US/09/981,353
; CURRENT FILING DATE: 2001-10-11
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: Perl Program
; SEQ ID NO 188
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020160382A1 522433CDI
; US-09-981-353-188
```

```
Query Match      87.1%; Score 27; DB 9; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      59 RYEDL 63

RESULT 21
; Sequence 6, Application US/09276600
; Patient No. US20010010908A1
; GENERAL INFORMATION:
; APPLICANT: Patricia Billing-Medel
; APPLICANT: Maurice Cohen
; APPLICANT: Tracey L. Colpitts
; APPLICANT: Julian Gordon
; APPLICANT: Edward N. Granados
; APPLICANT: John C. Russell
; APPLICANT: Stephen D. Stroupe
; TITLE OF INVENTION: Reagents and Methods useful for
; TITLE OF INVENTION: Detecting Disease of the Prostate
; FILE REFERENCE: 6397.US.01
; CURRENT APPLICATION NUMBER: US/09/276,600
; CURRENT FILING DATE: 1999-03-25
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 308
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-276-600-6

Query Match      87.1%; Score 27; DB 10; Length 308;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2 RYEDL 6
        |||||
Db      59 RYEDL 63

RESULT 22
; Sequence 14, Application US/09966546
; Patient No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20020168716A1 Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; CURRENT FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-966-546-14

Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 RYED 5
        |||||
```

Db 262 VYED 266

RESULT 23

```
US-09-966-545-14
; Sequence 14, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20020172999A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-343-14
```

```
Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYED 5

Db 262 VYED 266

RESULT 24

```
US-09-965-212-14
; Sequence 14, Application US/09965212
; Publication No. US20030003462A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A.
; TITLE OF INVENTION: No. US20030003462A1 Human Proteins and Polynucleotides Encoding
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/965,212
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: US/09/544,511
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: USN 60/128,514
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 14
; LENGTH: 316
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-965-212-14
```

```
Query Match      87.1%; Score 27; DB 9; Length 316;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYED 5

Db 262 VYED 266

RESULT 25

```
US-09-927-602-4
; Sequence 4, Application US/09927602
; Patent No. US20020061562A1
```

GENERAL INFORMATION:

```
APPLICANT: Fukuda, Michiko N.
APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-4
```

```
Query Match      87.1%; Score 27; DB 10; Length 390;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYEDL 6

Db 272 VYEDL 277

RESULT 26

```
US-09-927-602-2
; Sequence 2, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-927-602-2
```

```
Query Match      87.1%; Score 27; DB 10; Length 395;
Best Local Similarity 83.3%; Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 VYEDL 6

Db 271 VYEDL 276

RESULT 27

```
US-09-927-602-3
; Sequence 3, Application US/09927602
; Patent No. US20020061562A1
; GENERAL INFORMATION:
; APPLICANT: Fukuda, Michiko N.
; APPLICANT: Akama, Tomoya O.
; TITLE OF INVENTION: Methods of Treating Macular Corneal
; FILE REFERENCE: P-LJ 4852
; CURRENT APPLICATION NUMBER: US/09/927,602
; CURRENT FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: US 09/638,211
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 38
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
; NAME/KEY: VARIANT
; LOCATION: (1)...(395)
; OTHER INFORMATION: Xaa = any amino acid
; US-09-927-602-3

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 395;
Pred. No. 1.6e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYEDL 6
Db 271 RYEDL 276

RESULT 28
US-09-815-242-12541
; Sequence 12541 Application US/09815242
; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12541
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12541

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 396;
Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 119 RYEDL 123

RESULT 29
US-09-815-242-12900
; Sequence 12900 Application US/09815242
```

```

; Patent No. US2002061569A1
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12900
; LENGTH: 396
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
; US-09-815-242-12900

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 396;
Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6
Db 119 RYEDL 123

RESULT 30
US-09-741-669-300
; Sequence 300 Application US/09741669
; Patent No. US2002022718A1
; GENERAL INFORMATION:
; APPLICANT: Forsyth, R. Allyn
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; TITLE OF INVENTION: Genes identified as required for
; FILE REFERENCE: ELITRA 009A
; CURRENT APPLICATION NUMBER: US/09/741,669
; CURRENT FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: US 60/173005
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 481
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 300
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Escherichia coli
; US-09-741-669-300

Query Match
Best Local Similarity 87.1%; Score 27; DB 10; Length 439;
Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 2 RYEDL 6
11111
DB 112 RYEDL 116

RESULT 31

US-09-815-242-10284
; Sequence 10284, Application US/09815242
; Patent No. US20020061569A1
; GENERAL INFORMATION:

APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari L.
APPLICANT: Zvackind, Judith W.
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John D.
APPLICANT: Carr, Grant J.
APPLICANT: Yamamoto, Robert T.

APPLICANT: Xu, H. Howard
TITLE OF INVENTION: Identification of Essential Genes in
FILE REFERENCE: ELTRA 011A

CURRENT APPLICATION NUMBER: US/09/815,242

PRIOR FILING DATE: 2001-03-21

PRIOR APPLICATION NUMBER: 60/191,078

PRIOR FILING DATE: 2000-03-21

PRIOR APPLICATION NUMBER: 60/206,848

PRIOR FILING DATE: 2000-05-23

PRIOR APPLICATION NUMBER: 60/207,727

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 60/242,578

PRIOR FILING DATE: 2000-10-23

PRIOR APPLICATION NUMBER: 60/253,625

PRIOR FILING DATE: 2000-11-27

PRIOR APPLICATION NUMBER: 60/257,931

PRIOR FILING DATE: 2000-12-23

PRIOR APPLICATION NUMBER: 60/269,308

PRIOR FILING DATE: 2001-02-16

NUMBER OF SEQ ID NOS: 14110

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 10284

LENGTH: 439

TYPE: PRT

ORGANISM: Escherichia coli

US-09-815-242-10284

Query Match 87.1%; Score 27; DB 10; Length 439;

Best Local Similarity 100.0%; Pred. No. 1.8e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6

11111

DB 112 RYEDL 116

RESULT 32

US-09-907-509-2
; Sequence 2, Application US/09907509
; Patent No. US20020090705A1
; GENERAL INFORMATION:

APPLICANT: Meyers, Rachel
TITLE OF INVENTION: 62088, A NOVEL HUMAN NUCLEOSIDE
FILE REFERENCE: MNT-177
CURRENT APPLICATION NUMBER: US/09/907,509

PRIOR FILING DATE: 2001-07-16

PRIOR APPLICATION NUMBER: 60/216385

NUMBER OF SEQ ID NOS: 3

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 2

LENGTH: 604

TYPE: PRT

ORGANISM: Homo sapiens

US-09-907-509-2

Query Match 87.1%; Score 27; DB 10; Length 604;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RYEDL 6

11111

DB 323 RYEDL 327

RESULT 33

US-09-771-467C-4
; Sequence 4, Application US/09771467C
; Patent No. US20020177549A1
; GENERAL INFORMATION:

APPLICANT: Luo, Yuling
Xiomel, Xu
TITLE OF INVENTION: Semaphorin K1 Polypeptides
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: SCIENCE & TECHNOLOGY LAW GROUP
STREET: 75 DENISE DRIVE
CITY: HILLSBOROUGH
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94010

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/771,467C
FILING DATE: 26-Jan-2001
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: OSMAN, RICHARD A
REGISTRATION NUMBER: 36,627
REFERENCE/DOCKET NUMBER: EXEL98-001

TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 343-4341
TELEFAX: (650) 343-4342

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 606 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 4:

US-09-771-467C-4

Query Match 87.1%; Score 27; DB 9; Length 606;

Best Local Similarity 100.0%; Pred. No. 2.5e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RYED 5

11111

DB 415 RYED 419

RESULT 34

US-09-808-602-59
; Sequence 59, Application US/09808602
; Patent No. US20020135115A1
; GENERAL INFORMATION:

APPLICANT: Vernet, Corine A
APPLICANT: Fernandes, Elma
APPLICANT: Shinkets, Richard A

APPLICANT: Herrman, John L

APPLICANT: Majumder, Kumud

APPLICANT: Mishra, Vishnu

APPLICANT: Mezes, Peter S

```

; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 59
; LENGTH: 693
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-59
```

```

Query Match      87.1%; Score 27; DB 9; Length 693;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VRYED 5
        |||||
DB      254 VRYED 258
```

```

RESULT 35
US-09-808-602-60
; Sequence 60, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 60
; LENGTH: 773
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-60
```

```

Query Match      87.1%; Score 27; DB 9; Length 773;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VRYED 5
        |||||
DB      329 VRYED 333
```

```

RESULT 36
US-09-808-602-4
; Sequence 4, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
```

```

; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; TITLE OF INVENTION: No. US20020155115A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 4
; LENGTH: 815
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-4
```

```

Query Match      87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VRYED 5
        |||||
DB      376 VRYED 380
```

```

RESULT 37
US-09-966-546-18
; Sequence 18, Application US/09966546
; Patent No. US20020168716A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20020168716A1el Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,546
; PRIOR FILING DATE: 2001-09-26
; PRIOR APPLICATION NUMBER: 09/544,511
; PRIOR FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 18
; LENGTH: 815
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-966-546-18
```

```

Query Match      87.1%; Score 27; DB 9; Length 815;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 VRYED 5
        |||||
DB      376 VRYED 380
```

```

RESULT 38
US-09-966-545-18
; Sequence 18, Application US/09966545
; Patent No. US20020172999A1
; GENERAL INFORMATION:
; APPLICANT: Fernandes, Elma
; APPLICANT: Vernet, Corine
; APPLICANT: Shinkets, Richard A
; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encod
; FILE REFERENCE: Cura-46 (15966-546)
; CURRENT APPLICATION NUMBER: US/09/966,545
```

;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 18
;; LENGTH: 815
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-545-18

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 815;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 376 VRYED 380

RESULT 39
US-09-965-212-18
;; Sequence 18, Application US/09965212
;; Publication No. US20030003462A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20030003462A1el Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/965,212
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: US/09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: USSN 60/128,514
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 18
;; LENGTH: 815
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-212-18

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 815;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 376 VRYED 380

RESULT 40
US-09-966-546-16
;; Sequence 16, Application US/0996546
;; Patent No. US2002018716A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US2002018716A1el Human Proteins and Polynucleotides Encoding
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,546
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842

;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-546-16

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 842;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 403 VRYED 407

RESULT 41
US-09-966-545-16
;; Sequence 16, Application US/09966545
;; Patent No. US20020172999A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20020172999A1el Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/966,545
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: 09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-966-545-16

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 842;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYED 5
|||||
DB 403 VRYED 407

RESULT 42
US-09-965-212-16
;; Sequence 16, Application US/09965212
;; Publication No. US20030003462A1
;; GENERAL INFORMATION:
;; APPLICANT: Fernandes, Elma
;; APPLICANT: Vernet, Corine
;; APPLICANT: Shinkets, Richard A.
;; TITLE OF INVENTION: No. US20030003462A1el Human Proteins and Polynucleotides Encod
;; FILE REFERENCE: Cura-46 (15966-546)
;; CURRENT APPLICATION NUMBER: US/09/965,212
;; CURRENT FILING DATE: 2001-09-26
;; PRIOR APPLICATION NUMBER: US/09/544,511
;; PRIOR FILING DATE: 2000-04-06
;; PRIOR APPLICATION NUMBER: USSN 60/128,514
;; PRIOR FILING DATE: 1999-04-09
;; NUMBER OF SEQ ID NOS: 57
;; SOFTWARE: Patentln Ver. 2.0
;; SEQ ID NO 16
;; LENGTH: 842
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-965-212-16

Query Match
Best Local Similarity 100.0%; Score 27; DB 9; Length 842;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 403 VRYED 407

RESULT 43

US-09-808-602-58
; Sequence 58, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 58
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-58

Query Match 87.1%; Score 27; DB 9; Length 850;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 406 VRYED 410

RESULT 44

US-09-808-602-73
; Sequence 73, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
; APPLICANT: Vernet, Corine A
; APPLICANT: Fernandes, Elma
; APPLICANT: Shinkets, Richard A
; APPLICANT: Herrman, John L
; APPLICANT: Majumder, Kumud
; APPLICANT: Mishra, Vishnu
; APPLICANT: Mezes, Peter S
; APPLICANT: MacDougall, John
; TITLE OF INVENTION: No. US20020155115A1 Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 15966-697 CIP
; CURRENT APPLICATION NUMBER: US/09/808,602
; PRIOR FILING DATE: 2001-03-14
; PRIOR APPLICATION NUMBER: 09/800,198
; PRIOR FILING DATE: 2001-03-05
; PRIOR APPLICATION NUMBER: 60/186,596
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 73
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-808-602-73

Query Match 87.1%; Score 27; DB 9; Length 983;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 539 VRYED 543

RESULT 45

US-10-013-136-2
; Sequence 2, Application US/10013136
; Publication No. US20030023067A1
; GENERAL INFORMATION:
; APPLICANT: Conklin, Darrell
; APPLICANT: Elsworth, Jeff L.
; TITLE OF INVENTION: FOLLISTATIN RELATED PROTEIN ZFSTA2
; FILE REFERENCE: 98-50
; CURRENT APPLICATION NUMBER: US/10/013,136
; PRIOR FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/412,554
; PRIOR FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 983
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-013-136-2

Query Match 87.1%; Score 27; DB 9; Length 983;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5
DB 539 VRYED 543

RESULT 46

US-09-782-714-2
; Sequence 2, Application US/09782714
; Patent No. US20020034808A1
; GENERAL INFORMATION:
; APPLICANT: Darst, Seth A
; APPLICANT: Zhang, Gongyi
; APPLICANT: Campbell, Elizabeth
; APPLICANT: Minakin, Leonid
; APPLICANT: Severinov, Konstantin
; TITLE OF INVENTION: A CRYSTAL OF BACTERIAL CORE RNA POLYMERASE AND METHODS
; FILE REFERENCE: 600-1-258
; CURRENT APPLICATION NUMBER: US/09/782,714
; PRIOR FILING DATE: 2001-02-13
; PRIOR APPLICATION NUMBER: 09/396,651
; PRIOR FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1119
; TYPE: PRT
; ORGANISM: Thermus aquaticus
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (695)..(696)
; OTHER INFORMATION: Any amino acids can be at these two positions.
US-09-782-714-2

Query Match 87.1%; Score 27; DB 10; Length 1119;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 VRYED 5

Db 613 VRYED 617

RESULT 47
US-09-864-761-34609
Sequence 34609, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Acmica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
CURRENT FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/652,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263,6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
SEQ ID NO 34609
LENGTH: 45
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC007383.2
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.8
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 9.8
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.7
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.2
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.4
OTHER INFORMATION: EXPRESSED IN BF474, SIGNAL = 4.1
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
OTHER INFORMATION: SWISSPROT HIT: P28531, EVALU 9.00e-20

OTHER INFORMATION: EST_HUMAN HIT: AW614813.1, EVALU 3.00e-19
US-09-864-761-34609

Query Match 83.9%; Score 26; DB 10; Length 45;
Best Local Similarity 66.7%; Pred. No. 30;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 26 VRYDDI 31

RESULT 48
US-09-764-878-127

Sequence 127, Application US/09764878
Patent No. US20020090615A1
GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PAL21

CURRENT APPLICATION NUMBER: US/09/764,878

CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 428

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 127

LENGTH: 144

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (25)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-878-127

Query Match 83.9%; Score 26; DB 10; Length 144;
Best Local Similarity 66.7%; Pred. No. 97;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 71 VRYDDI 76

RESULT 49

US-09-764-860-461

Sequence 461, Application US/09764860

Patent No. US20020094953A1

GENERAL INFORMATION:

APPLICANT: Rosen et al.

TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

FILE REFERENCE: PC008

CURRENT APPLICATION NUMBER: US/09/764,860

CURRENT FILING DATE: 2001-01-17

Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 1198

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 461

LENGTH: 144

TYPE: PRT

ORGANISM: Homo sapiens

FEATURE:

NAME/KEY: SITE

LOCATION: (25)

OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-764-860-461

Query Match 83.9%; Score 26; DB 10; Length 144;
Best Local Similarity 66.7%; Pred. No. 97;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 VRYEDL 6
Db 1 VRYDDI 6

Db 71 VRYDDI 76

RESULT 50
 US-09-764-868-919
 ; Sequence 919, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: Patentln Ver. 2.0
 ; SEQ ID NO 919
 ; LENGTH: 155
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-764-868-919

Query Match 83.9%; Score 26; DB 9; Length 155;
 Best Local Similarity 66.7%; Pred. No. 1e+02;
 Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VRYDDL 6
 :||:|
 Db 89 IRYKDL 94

Search completed: February 20, 2003, 13:40:58
 Job time : 12.7429 secs

This Page Blank (uspto)